ARIZONA STATE UNIVERSITY

The Inevitable Corruption of Indicators and Educators Through High-Stakes Testing

by

Sharon L. Nichols
Assistant Professor
University of Texas at San Antonio

and

David C. Berliner
Regents' Professor
Arizona State University

Education Policy Research Unit (EPRU)
Education Policy Studies Laboratory
College of Education
Division of Educational Leadership and Policy Studies
Box 872411
Arizona State University
Tempe, AZ 85287-2411

March 2005

EPSL | EDUCATION POLICY STUDIES LABORATORY
Education Policy Research Unit

EPSL-0503-101-EPRU
http://edpolicylab.org

Education Policy Studies Laboratory
Division of Educational Leadership and Policy Studies
College of Education, Arizona State University
P.O. Box 872411, Tempe, AZ 85287-2411
Telephone: (480) 965-1886
Fax: (480) 965-0303
E-mail: epsl@asu.edu
http://edpolicylab.org

This research was made possible by a grant from the Great Lakes Center for Education Research and Practice.

Table of Contents

The Inevitable Corruption of Indicators and Educators Through High-Stakes Testing

Sharon L. Nichols

University of Texas at San Antonio

and

David C. Berliner

Arizona State University

Executive Summary

This research provides lengthy proof of a principle of social science known as Campbell's law: "The more any quantitative social indicator is used for social decision-making, the more subject it will be to corruption pressures and the more apt it will be to distort and corrupt the social processes it is intended to monitor."[1] Applying this principle, this study finds that the over-reliance on high-stakes testing has serious negative repercussions that are present at every level of the public school system.

Standardized-test scores and other variables used for judging the performance of school districts have become corruptible indicators because of the high stakes attached to them. These include future employability of teachers and administrators, bonus pay for school personnel, promotion/non-promotion of a student to a higher grade, achievement/non-achievement of a high school degree, reconstitution of a school, and losses or gains in federal and state funding received by a school or school district.

Evidence of Campbell's law at work was found in hundreds of news stories across America, and almost all were written in the last few years. The stories were gathered using LexisNexis, Inbox Robot, Google News Alerts, *The New York Times*, and Ed Week Online. In addition to news stories, traditional research studies, and stories told by educators about the effects of high-stakes testing are also part of the data. The data fell into 10 categories. Taken together these data reveal a striking picture of the corrupting effects of high-stakes testing:

1. **Administrator and Teacher Cheating:** In Texas, an administrator gave students who performed poorly on past standardized tests incorrect ID numbers to ensure their scores would not count toward the district average.
2. **Student Cheating:** Nearly half of 2,000 students in an online Gallop poll admitted they have cheated at least once on an exam or test. Some students said they were surprised that the percentage was not higher.
3. **Exclusion of Low-Performance Students From Testing:** In Tampa, a student who had a low GPA and failed portions of the state's standardized exam received a letter from the school encouraging him to drop out even though he was eligible to stay, take more courses to bring up his GPA, and retake the standardized exam.
4. **Misrepresentation of Student Dropouts:** In New York, thousands of students were counseled to leave high school and to try their hand at high school equivalency programs. Students who enrolled in equivalency programs did not count as dropouts and did not have to pass the Regents' exams necessary for a high-school diploma.

5. **Teaching to the Test:** Teachers are forced to cut creative elements of their curriculum like art, creative writing, and hands-on activities to prepare students for the standardized tests. In some cases, when standardized tests focus on math and reading skills, teachers abandon traditional subjects like social studies and science to drill students on test-taking skills.
6. **Narrowing the Curriculum:** In Florida, a fourth-grade teacher showed her students how to navigate through a 45-minute essay portion of the state's standardized exam. The lesson was helpful for the test, but detrimental to emerging writers because it diluted their creativity and forced them to write in a rigid format.
7. **Conflicting Accountability Ratings:** In North Carolina, 32 schools rated excellent by the state failed to make federally mandated progress.
8. **Questions about the Meaning of Proficiency:** After raising achievement benchmarks, Maine considered lowering them over concerns that higher standards will hurt the state when it comes to No Child Left Behind.
9. **Declining Teacher Morale:** A South Carolina sixth-grade teacher felt the pressure of standardized tests because she said her career was in the hands of 12-year-old students.
10. **Score Reporting Errors:** Harcourt Educational Measurement was hit with a $1.1 million fine for incorrectly grading 440,000 tests in California, accounting for more than 10 percent of the tests taken in the state that year.

High-stakes tests cannot be trusted – they are corrupted and distorted. To avoid exhaustive investigations into these tests that turn educators into police, this research supports building a new indicator system that is not subject to the distortions of high-stakes testing.

Introduction

The United States faces a severe crisis. Because the dangers do not seem imminent, the few individuals and organizations alerting politicians and federal agencies to the crisis are generally unheeded. This crisis concerns the corruption of what is arguably America's greatest invention—its public schools.

This research joins with others in documenting the damage to education caused by overreliance on high-stakes testing. Our documentation suggests that the incidence of negative events associated with high-stakes testing is so great, corruption is inevitable and widespread. As will be made clear, below, public education is presenting serious and harmful symptoms. Unlike other critics of the high-stakes testing movement, however, we demonstrate that a powerful social science law explains the etiology of the problems we document. Ignorance of this law endangers the health of our schools.

Criticisms of Testing

Concerns about the negative effects associated with testing are certainly not new, as demonstrated by comments from the Department of Education of the state of New York:

> *It is an evil for a well-taught and well-trained student to fail in an examination.*
>
> *It is an evil for an unqualified student, through some inefficiency of the test, to obtain credit in an examination.*
>
> *It is a great and more serious evil, by too frequent and too numerous examinations, so to magnify their importance that students come to regard them not as a means in education but as the final purpose, the ultimate goal.*

> *It is a very great and more serious evil to sacrifice systematic instruction and a comprehensive view of the subject for the scrappy and unrelated knowledge gained by students who are persistently drilled in the mere answering of questions issued by the Education Department or other governing bodies.*

This Department of Education raises issues about the reliability and validity of its tests, as every testing agency should. But they also are concerned about over-testing and about how testing programs can mislead students (and by implication—parents and politicians) into thinking test scores are indicators of a sound education. This Department of Education also expresses its worry about how testing can distort the educational system by narrowing the curriculum. The enlightened bureaucrats who wrote this report to the legislature were warning the state's politicians that it is possible, with the best of intents, for testing programs to corrupt the educational process. The archaic language in their report is better understood if the date of the report is known. It was written in 1906.[2]

Another warning about corruption and distortion from high-stakes tests surfaced when a plan to pay teachers on the basis of their students' scores was offered, making indicators of student achievement very high-stakes for teachers. A schoolmaster noted that under these conditions

> *… a teacher knows that his whole professional status depends on the results he produces and he is really turned into a machine for producing these results; that is, I think, unaccompanied by any substantial gain to the whole cause of education."*

This document is available on the Education Policy Studies Laboratory website at:
http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc

This concern about testing students to judge a teachers' worth first surfaced in the year 1887,[3] but it is as fresh as this year's headlines about a new pay-for-performance plan in Denver, Colorado.[4]

These two criticisms of what we now call high-stakes testing were made before modern testing systems had gained admiration for their beneficial effects on education, and when a crisis seemed quite far away. Therefore the minor worries of an individual here and there, over the last century, were easily set aside. But today, high-stakes testing in the United States is more widespread than ever before, and our nation apparently relies on ability and achievement testing more than any other nation for making important decisions about individuals and schools. We live in an era and in a nation where there is strong support for public policies that use test results to compel changes in the behavior of students, teachers, and school administrators. Our President, politicians from both parties, and many citizens believe that education can best be improved by attaching consequences (that is, attaching high stakes) to tests. The tests are seen by some as the perfect policy mechanism because they are both *effectors* and *detectors*—they are intended to *effect* or cause change in the system, and then *detect* whether changes in the system actually occur. The federal No Child Left Behind (NCLB) Act of 2001 sanctifies this approach to school change.

As might be expected from increased reliance on tests to make important decisions, there now exist a far greater number of critics among whom are highly respected scholars, social critics and scholarly organizations. Among these are: Gerald Bracey,[5] Robert Brennan,[6] Center for the Study of Testing, Evaluation and Educational Policy,[7] Fairtest [8] Robert Linn,[9] Jay Heubert and Robert Hauser writing for the National

This document is available on the Education Policy Studies Laboratory website at:
http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc

Research Council/National Academy of Sciences,[10] Lyle Jones,[11] Alfie Kohn,[12] National Board on Educational Testing and Public Policy,[13] Susan Ohanian,[14] Gary Orfield and Mindy Kornhaber of the Harvard University Civil Rights Project,[15] and Stephen Raudenbush.[16]

The numbers of individuals speaking out, along with the quality and the passion of their arguments, suggest that there is a crisis. But few of these critics have used the power of Donald Campbell's well-established social science law to support their arguments.

Campbell's law has two parts, one of which is concerned with the validity of the indicators we use, and one of which is concerned with the organizations and the people that work with indicators when they take on exceptional value. Campbell states, "The more any quantitative social indicator is used for social decision-making, the more subject it will be to corruption pressures and the more apt it will be to distort and corrupt the social processes it is intended to monitor."[17] George Madaus[18] has pointed out that Campbell has given the social sciences a version of the Heisenberg uncertainty principle. That principle, concerned with measuring the position and velocity of objects, informed physicists that if they measure one of these conditions they could not accurately measure the other at the same time. Madaus' version of the uncertainty principle with regard to Campbell's Law states that if you use high-stakes tests to assess students, teachers, or schools, the corruptions and distortions that inevitably appear compromise the construct validity of the test. As the

As the stakes associated with a test go up, so does the uncertainty about the meaning of a score on the test.

This document is available on the Education Policy Studies Laboratory website at:
http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc

stakes associated with a test go up, so does the uncertainty about the meaning of a score on the test. That is, in high-stakes testing environments, the greater the pressure to do well on the tests the more likely is the meaning of the score obtained by students or schools uninterpretable.

Serious, life-altering decisions are made on the basis of high-stakes tests, such as promotion to a higher grade or retention in grade. Tests can determine who will receive a high-school degree, and who will not. Tests scores can determine if a school will be reconstituted, with job losses for teachers and administrators when scores do not improve or cash bonuses when scores do improve. Thus test-givers should be certain that the construct measured by those who take tests with serious consequences attached to them is the construct that was intended. Too much uncertainty about the meaning of a score on a test would be psychometrically, morally, and (sometimes) legally inappropriate. It also violates the standards that professionals and their professional organizations have agreed to use when constructing and administering tests.[19]

The remainder of this paper provides an excess of examples of Campbell's law and Madaus' principle illustrating their ubiquity in commerce, government, and education. In a wide range of human endeavors, problems are frequently noted when the endeavors are judged by indicators to which serious consequences are attached. These problems are as likely to be associated with track and field events, factory production, or police reports as they are end-of-semester tests. Wherever we seem to look, when high-stakes are attached to indicators, what follows is the corruption and distortion of the indicators and the people who employ them. Examples illustrating the general case of

This document is available on the Education Policy Studies Laboratory website at:
http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc

Campbell's law are provided next, after which we provide specific examples of Campbell's law in our educational system.

Corrupting the Indicators and the People in the World Outside of Education

Corruption in Business

In the world of business the corruption and distortion of indicators and people when the stakes are high is well known. This raises the issue, of course, about why anyone would want to bring a failed incentive system to education. For example, an article in the *Quarterly Journal of Economics*[20] notes:

> At the H.J. Heinz Company, division managers received bonuses only if earnings increased from the prior year. The managers delivered consistent earnings growth by manipulating the timing of shipments to customers and by prepaying for services not yet received, both at some cost to the firm. At Dun & Bradstreet, salespeople earned no commission unless the customer bought a larger subscription to the firm's credit-report services than in the previous year. In 1989, the company faced millions of dollars in lawsuits following charges that its salespeople deceived customers into buying larger subscriptions by fraudulently overstating their historical usage. In 1992, Sears abolished the commission plan in its auto-repair shops, which paid mechanics based on the profits from repairs authorized by customers. Mechanics misled customers into authorizing

This document is available on the Education Policy Studies Laboratory website at:
http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc

unnecessary repairs, leading California officials to prepare to close Sears' auto-repair business statewide.

Events like those that occurred at Heinz seem also to have occurred at other large corporations, including Qwest and Enron. At Qwest "Prosecutors alleged [that] four executives, under heavy pressure to meet revenue goals, used the deal with [an Arizona client] to book revenue the company did not see until six months later, and then lied to accountants and investigators about it."[21] At Enron "Prosecutors argue in court papers that former Enron Corp. chief executives Jeffrey K. Skilling and Kenneth L. Lay should stand trial together because they engaged in a "single overarching conspiracy to enrich themselves by inflating the company's stock price."[22] Not only did Arthur Anderson, Enron's auditing firm, not watch out for the public's interest, they too were corrupted by the stakes involved. They hid Enron's improprieties, earning a lot of money for their company, though ultimately being forced out of business completely by their corrupted accounting practices whose origins were in simple greed. They were not alone. In the Enron case, two large banks sworn to protect the public interest and to abide by the federal regulations for the banking industry failed do so. In that scandal J.P. Morgan and CitiCorp were found to have readily violated the public trust. They were corrupted, engaging in unacceptable banking practices because the stakes were high. They were asked to pay the government fines of $147.5 million and $132.5 million, respectively. Lehman Brothers securities and three British bankers were also indicted for fraud. All of these institutions aspire to earn integrity but were easily corrupted by Enron's schemes. Those schemes were simply a version of the fact that when stakes get high the indicators

This document is available on the Education Policy Studies Laboratory website at:
http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc

used (income, assets, accounts receivable, reserves, accounts payable, outstanding liabilities) all become corrupted, as do the people who work in those firms.

In business, high-stakes are associated with increasing stock prices, something easy to accomplish by manipulating the perception that corporate profits appear to be on the rise. For executives and privileged stockholders a great deal of money can be made if profits look to be going up. To promote that perception, Qwest and Enron recorded profits on sales not yet made. This is a patently illegal manipulation of the stock price, which is the "score" by which the general public and accountancy firms judge how well corporations are doing. While business has stock price as its major indictor, education has achievement test scores as its major indicator, and therefore, education is subject to the same corrupting forces. The general rule is that any indicator is subject to corruption when the stakes become too high. Corruption of the Qwest and Enron personnel was widespread. The court records provide an ugly picture for the nation to contemplate, although the scandal was unsurprising given the ubiquity of Campbell's law.

Most recently Southern California Edison acted in accordance with Campbell's law.[23] Edison admitted to falsifying workplace safety data and might have also suppressed reports of on-the-job injuries over a period of seven years in order to win performance bonuses from the state. Edison's falsifying and hiding of medical records left an attorney for the Public Utility Commission "flabbergasted." The attorney went on to say what Donald Campbell might have said "What this appears to be is an incentive…. to underreport injuries. That's what happened here." This sad state of affairs should have been predicted, especially since Edison had previously admitted to falsifying data by having both employees and managers rig customer satisfaction surveys to win millions in

This document is available on the Education Policy Studies Laboratory website at:
http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc

bonuses from the state. This example shows that when the stakes are high, oversight agencies need to do a better job of checking the validity of the indicators that are used to determine bonuses and negative sanctions

A last instance of this kind of corruption in business is as funny as it is sad. It is about the business side of education, the sports programs at universities. Like any other business enterprise, sports programs in universities understand that they have an obligation to make money and bring prestige to their institution through exceptional athletic teams. Because of the stakes involved beyond merely winning or losing athletic contests, the National Collegiate Athletic Association (NCAA) has had to "police" universities, as the Security and Exchange Commission (SEC), or Federal Aviation Authority (FAA), or Food and Drug Administration (FDA) is to do for securities firms, airline, and pharmaceutical companies. But most policing activities of this kind are under-funded and understaffed. Thus the agencies with oversight ordinarily put a great deal of "trust" in the organizations they are required to watch over. Hence, they do much less policing than might be warranted, given the predictive power of Campbell's law and its ubiquity. As a consequence of the stakes involved almost all institutions with competitive sports teams have been warned, fined, or placed on probation over the years for violating the NCAA rules. But warnings and fines have also been the lot of Lehman Brothers and Merrill Lynch in the securities area, Alaska Air and America West among airlines, and Merck and Squibb among pharmaceutical houses, just to name only a few of the many examples that could be noted as having violated guidelines for appropriate behavior.

This document is available on the Education Policy Studies Laboratory website at:
http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc

Focusing back on intercollegiate sports, we can imagine that the consequences would be serious, say, for a basketball team with a losing record. Thus it is likely that indictors and personnel associated with college basketball could become corrupted. This is exactly what happened recently at the University of Georgia where basketball is big business.[24] Assistant Coach Jim Harrick Jr. taught *Coaching Principles and Strategies of Basketball*. It was important for the basketball students he coached to have high grade-point averages in order to play, and so many took this elective course. The 20-question final exam contained items that make this example of Campbell's law funnier than it really should be. Here are four examples:

1. How many goals are on a basketball court?

a. 1 b. 2 c. 3 d. 4

2. How many players are allowed to play at one time on any one team in a regulation game?

a. 2 b. 3 c. 4 d. 5

5. How many halves are in a college basketball game?

a. 1 b. 2 c. 3 d. 4

And, apparently hoping to get away from simple multiple-choice measures of assessment, Coach Harrick introduced a "performance" item:

17. Diagram the half-court line.

In this example we see how the stakes involved corrupted the indicators used for assessing academic competence, and, of course, the coach and the students themselves.

This document is available on the Education Policy Studies Laboratory website at:
http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc

The high stakes associated with collegiate sports corrupts the whole system, giving the nation recurring reports of violations of recruiting procedures through the provision of gifts and promises to talented young athletes, the provision of easy courses and grading for athletes at the college, and by offering leniency when athletes are charged with crimes. We hear also about the harshness of treatment that injured athletes receive, such as the loss of scholarship, when their contributions to salient indicators are compromised. These all-to-common events occur only when winning records in a sport take on special significance as indicators of a university's "greatness." But as will be noted in the next section of this report, it is no different in education. Test scores, rather than basketball scores, have become the indicators of "greatness" for a school or a district. And so we see many instances of the academically talented students having privileges while the less talented are treated as throwaways, just as is in athletics.

Winning records in a sport take on special significance as indicators of a university's "greatness." It is no different in education. Test scores have become the indicators of "greatness" for a school or a district. And so we see many instances of the academically talented students having privileges while the less talented are treated as throwaways, just as is in athletics.

Corruption in Medicine

Physicians, just like Heinz, Qwest, and Enron, are also judged by means of certain indicators.[25] But if the indicators by which physicians are judged take on considerable importance, then Campbell's law predicts that the indicator and the physicians will both

This document is available on the Education Policy Studies Laboratory website at:
http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc

be corrupted. This appears to be the case, thus giving rise to more expensive medical costs for Americans.

For example, physicians are evaluated on the two types of errors that they make. They can diagnose a healthy person as being sick (Type I error) or they can diagnose a sick person as being well (Type II error). Obviously there are serious consequences associated with either type of error, but the consequences for making a Type II error are much more severe. Type II errors can result in expensive litigation and loss of reputation. Since the numbers of diagnoses of genuinely sick people as healthy people must be minimized by physicians, it should not be surprising to find out that physicians have found ways to avoid making Type II errors. So much weight is on that indicator, the number of mistaken diagnosis, that physicians find a way to rarely get caught doing that.

They avoid Type II errors by over-diagnosing illness among their healthy patients. That is, physicians purposefully inflate their Type I errors to deflate their Type II errors! The result of this is that many people believe they are sick when they are not. Treatments such as invasive surgery and pharmaceuticals are prescribed though they are not needed. In addition, to also keep Type II errors down, physicians require more testing for illnesses that are not likely. Of course the cost of medicine is driven up by these methods to avoid the dreaded Type II error. As Campbell's law predicts, the more importance that an indicator takes on, the more likely *it*, and the people who depend on it, will be corrupted.

Other examples of corruption exist in medicine, often around money, historically the indicator around which corruption so often takes place. Periodically we hear of drug companies providing monetary kick backs or free trips to conventions to physicians who

This document is available on the Education Policy Studies Laboratory website at:
http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc

use their drugs; the payments of kick backs for referrals between physicians; bonus' for physicians in HMO's who do not use laboratory tests as frequently; and so forth. Physicians, like teachers, are supposed to be paragons of virtue, but Campbell's law often overrides community expectations.

Corruption in Sports and Academe

Corruption shows up quite clearly when a job has many components, such as that of being an outfielder for a baseball team, or a professor at a university. The contracts one receives to play ball or to profess do not specify how many homeruns have to be hit or how many publications need to be completed per year. But it soon becomes clear to the baseball player and the professor that these indicators have the highest stakes attached to them. For the baseball player indicators used to assess fielding, throwing, bases-on-balls received, contributions to team morale, and so forth, are not nearly as important as the number of hits and especially the number of home runs that are hit. Home-runs pay very well and the other indicators of successful fulfillment of the job as a baseball player, including responsibility to the team and the fans, do not matter nearly as much. Therefore some players pursue hits and runs at the cost of other dimensions of their job, sometimes at the cost to their team, and sometimes at great cost to their health, as when they take steroids to increase strength and speed. The recent scandals about drug use in baseball, the Olympics, and other sporting events are about the corruption of indictors and athletes when stakes get too high. The corruption and distortion of the processes and the practices of the people who work in that field are easily predicted given the

This document is available on the Education Policy Studies Laboratory website at:
http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc

importance associated with the indicators that are used to judge their performance. Unfortunately the professoriate works in exactly the same way.

In the academic community corruption stems from the exaggerated emphasis on publications as *the* major criteria for determining the worth of a professor. Merit pay, ability to move from one institution to another, tenure, and advancement to full professor, have all become dependent on publication. The other duties of the job such as teaching, meeting with students, participating in faculty governance, and service to the profession, often become secondary to being published. Campbell's law predicts the outcome: publications of little worth, publication of the same ideas or data in slightly different forms in different journals; a proliferation of journals of unknown quality in each discipline to accommodate the increased need for publications, and, indeed, an increased incidence of academic plagiarism.

Publications, like home runs in baseball, become the way in which quality is judged. As a consequence of associating high-stakes with this one part of a professor's job, the university has been corrupted. Its attention has been diverted from its many missions and allocated primarily to just one of its missions. The economists have a name for this. It is called multi-tasking, a form of gaming the system whereby those aspects of a job that are not compensated are sacrificed for those aspects of a job that are compensated. This kind of corruption is well documented in economics.[26] But the problems caused by multitasking are ignored in education where, more than ever before, the worth of a teacher or a school is determined by a single outcome: Student scores on achievement tests. Predictably, we can expect teachers and schools to multi-task, working hard to get test scores up and doing other parts of their jobs less well. As will be

This document is available on the Education Policy Studies Laboratory website at:
http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc

made clear, below, this is exactly what happens. It is a part of the corruption of the personnel that is predicted under Campbell's law.

Corruption in Politics

Other examples of the corruption of indicators and persons abound, especially in politics. Stakes have always been high in political elections, so we watch for corruption in vote tallying, and we have found lots of it over the years, as predicted by Campbell's law. New York's Boss Tweed and Chicago's first Mayor Daley each were famous, sometimes celebrated, for their corruption. The Florida voting irregularities in the 2000 presidential election raised the corruption issue in contemporary time, and none of these fears was allayed in the presidential election of 2004 where again it was suggested that the indicator—vote count—was corrupted. This is why, in part, most Americans currently express distrust in electronic voting machines. There seems to be an implicit knowledge by Americans of Campbell's law when it comes to voting, with many people understanding that as stakes get high so does the possibility of corruption. That same understanding is not as often applied to the educational system where high-stakes assessment of schools and students is similar to voting. Test scores, like votes, determine the winners and losers among schools and districts, and thus it is likely that attempts will be made to manipulate test scores, as is historically true of voting. The similarities between the two tell us that Campbell's law is operating in both realms.

Another political issue illustrating Campbell's law has to do with the census data, which for many years was characterized by low stakes. Thus most Americans paid the census no attention. But in recent decades the stakes associated with the census have

This document is available on the Education Policy Studies Laboratory website at:
http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc

gone up. Congressional districts are formed based on census counts, and large amounts of federal money are provided to states based on the numbers of certain classes of people identified by the census count. So now battles are waged over how the census counts are to be made. Currently, the census bureau is continuously challenged by the possibility of over- and under-counting people, resulting in (perhaps) unmerited monetary gains for some states and monetary losses for others. Since the stakes are very high, the possibility of corruption is great. It is proper to be vigilant about the census. High-stakes indicators attract corruption like honey attracts flies.

Corruption in Government Agencies

Many years ago Jerome Skolnick,[27] a leading sociologist of police work, described the corruption associated with an indicator used to determine police efficiency. When police departments were judged by clearance rates for crimes in their communities, they simply failed to report many crimes, or they down-graded the severity of crimes, and they even helped criminals they captured get lower sentences if the criminal would plea to many similar crimes so those crimes could get cleared off the books. It seemed clear to Donald Campbell[28] that Richard Nixon's heralded crackdown on crime resulted in the corruption of the crime rate statistics. The most immediate effect of having made the crime rate more important was the under-recording of crimes and their downgrading to less serious offences. Crime rates moved in the way the indicator needed to move by those for whom the indicator was high-stakes.

When the number of cases handled per unit time is the measure used to evaluate the efficiency in government offices, or a medical clinic, quick service is obtained but it

This document is available on the Education Policy Studies Laboratory website at:
http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc

is often not very good. And people are hurt, like foster children whose caseworkers have to make their visits short to handle their large case load. Judging efficiency and giving bonuses for meeting targets by measuring the time spent per client has led to angry clients of our nation's medical plans who feel cheated out of service. Worse, indicators concerned with efficiency have resulted in the death of too many patients and too many children in court-supervised child care facilities. In these cases the stakes attached to the indicator, the number of cases handled per hour or per day, were too high. The result was a corruption of the indicator and the corruption of the people associated with that indicator. Campbell's law once again demonstrates that it applies to a wide range of human activities.

The factory system during communism in the U.S.S.R. provides classic examples of the corruption of indicators.[29] Whatever measure the government used to judge factory output, those were the indicators that went up! If 100 tractors had to leave an assembly line each month, then 100 tractors left the line. But whether they worked or not or could even be found was a completely different matter. The problem, of course, was that the indicators and the factory managers were corrupted by the threats associated with not meeting quotas. Under threats—very high stakes for ones self and family—quotas were met, although real productivity in the U. S. S. R. never changed at all.

In our government, the Pentagon, naturally enough, places a heavy reliance on recruitment and retention figures in the various military services. Campbell's law predicts that some corruption of that indicator might then take place and The Los Angeles Times provided such evidence this year.[30] For over a decade the Pentagon was warned that they were accepting military recruits with criminal histories. They were also warned

This document is available on the Education Policy Studies Laboratory website at:
http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc

that the services were too lenient with violent individuals already in the ranks. This makes recruitment and retention figures look better but cannot be good for the nation's armed services. Unsurprisingly, one of the American's charged with murder in Afghanistan as well as the alleged ringleader of the Abu Graib prison scandal were repeated violent offenders *before* their service commitments.

Our safety has even been compromised by another branch of our government, which recently forgot Campbell's law. The relatively new Transportation Safety Authority (TSA), the guardians of our airports, needed to find and train employees quickly. Among the companies administering the contract to do this were Pearson and Thompson, both big publishing conglomerates with testing experience. They were to be judged on how fast they could get people to work, and it seems that their income would be enhanced if a competent work force were on-the-job quickly. As a result, the contractors found people and put them to work quickly. A federal inspector, however, found a few glitches in the process of getting those people to work. Questions from the final exam for airport screeners, designed to measure the important and intensive training that the screeners had completed, were a farce.[31] One of these questions was: "How do threats get on board an aircraft?" To this supposedly challenging question the answers candidates had to choose from were:

a. In carry-on bags;

b. In checked-in bags;

c. In another person's bag;

d. All of the above.

This document is available on the Education Policy Studies Laboratory website at:
http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc

As if an item like this is not frightening enough, Senator Shumer, who reviewed the TSA training program, said other items looked as if they were written by Jay Leno and not a testing company!

Senator Shumer, who reviewed the Transportation Safety Authority training program, said test questions looked as if they were written by Jay Leno and not a testing company.

Moreover, it appears that items like these might have been too difficult for some of the recruits! The inspectors of the program also found that 22 of the exam's 25 questions were repeats from previous exams and that some test-takers were given the questions in advance of the exam. The stakes corrupted the test-makers, the trainers of the workers, and the administrators of the training program. Indicators of successful training had no meaning as the rush to get employees on the job became the primary goal of management.

In the business side of education we saw examples of this same phenomena among the performance contractors of a few years back. That was a version of pay for performance as school managers, using test scores as outcomes. We include this example here as a business example, because it is about the inflation of scores that are important to the firms stock price, and thus no different from the Enron case. In these instances contractors taught the students the tests quite directly, and corrupted both the indicator and themselves. It appears that some of today's voucher and charter school operators are doing the same thing.

This document is available on the Education Policy Studies Laboratory website at:
http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc

Is Campbell's Law Ubiquitous?

From this brief introduction it does look like examples of Campbell's law are to be found everywhere. In a broad range of human endeavors, when the stakes get high, the indictors used to monitor events and processes become less trustworthy. The integrity of individuals who are associated with highly valued indicators also suffers. Perhaps, because it is so hard to maintain, this is why integrity has become such a valued characteristic of individuals and institutions.

Campbell's law appears well established, and it is supported by economic theory. Let us now examine how that law manifests itself in the systems of educational accountability now being developed across our nation.

Corrupting the Indicators and the People in Education

Methodology

The method used to find examples of Campbell's law in the field of education was a search of news stories throughout America. The search for stories was conducted from October 2003 through October 2004. We used three general approaches. The first approach was to conduct a one-time broad search of news stories using the LexisNexis database.[32] The purpose of this search was to identify relevant articles dating back to 1990 and to identify and refine the list of categories and key words we needed to guide subsequent searches.

This document is available on the Education Policy Studies Laboratory website at:
http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc

A second approach involved subscribing to news alert systems. We subscribed to two main systems that we felt would maximize the potential number and type of stories we wanted to analyze. These included Inbox Robot[33] and the Google news alert system.[34] Our subscription to these news alerts provided us with daily emails indicating the most recent news stories relevant to the key words associated with our search. A third approach involved periodic reviews of two primary news sources. These included the *New York Times* since it is has regular reporting of education stories and columns on education, and *Ed Week* online,[35] because it had a "daily archives" section that provided major education stories from news outlets in most states.

Our searches were deliberately broad. Our approach was to identify stories that would fit within the categories and searchable terms that we specified (achievement testing, high-stakes, teacher cheating, teacher morale, NCLB, and so forth). Therefore, the stories we present in abbreviated form, below, are not representative of the entire universe of educational stories available in the news media. Our method was simply to yield stories on topics that might reveal Campbell's law at work in education. In that we were successful. The number and type of stories we identified from across the nation and across a range of topics paint a striking picture of the corrupting effects of high-stakes testing on education practice.

What this search procedure cannot do is provide information about rates or frequencies. Therefore, these data do not report the national or local incidence of certain events. We catalog here only the occurrence of events as reported in the news stories we found. Given our inability to tap all news sources, the ambiguity of some of the stories we did find, and the fact that many of the incidents we report are often "hushed up" at the

This document is available on the Education Policy Studies Laboratory website at:
http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc

local levels, these data represent a vast underestimate of the incidence of the events themselves.

One other caveat about our search is important. Some of the stories we report are about *alleged* events. Not all of them necessarily turned out to be true or resulted in legal or administrative action. Unfortunately, follow-up stories to allegations are not always printed and if they are printed it is often months after the original incidents on which we report, making it hard to find and to match these reports with earlier ones. Thus we remind readers that we do attest to the validity of each newsworthy incident reported.

The interpretable stories have been grouped into ten tables that are discussed below. Nine of these tables are examples of Campbell's law in action. Table 10 emerged from our search and is only indirectly an example of Campbell's law—it deals with the inability of testing companies and state departments of education to handle the loads of data that high-stakes accountability systems produce. The table also documents the tragic effects for students and their families when test companies and state departments make mistakes. This attaches a human face to the errors. We have added that table to this report since it is related to the viability of the whole accountability system currently under construction in the United States. The table also demonstrates what can happen when a single indicator becomes highly valued for making decisions but is created under conditions where the people in charge of the indicator are trying to produce it at the least possible cost and the greatest possible profit. What follows, then, are ten subsections around the categories into which we grouped news reports. Each section is associated with a table in which abbreviated news reports (and other documents we discovered) are

This document is available on the Education Policy Studies Laboratory website at:
http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc

given. We start with the biggest of these tables, and the most illustrative of Campbell's law, the data on administrator and teacher cheating.

Administrator and Teacher Cheating

Table 1 presents numerous examples of educators that have been corrupted by programs of high-stakes testing. To be clear, we are not trying here to castigate administrators and teachers by identifying the unsavory practices in which some of them engage. About three million teachers and hundreds of thousands of administrators are practicing their profession honorably. The more important question is why have our politicians and lawmakers created a system that pressures people who we expect to be moral leaders of our youth? *That is what is most important to think about.* More specifically, the question to ask is why some of these professionals think the system they are in is so unfair to their schools that cheating, direct test preparation, and violating standardization procedures are justified? Another curious question that begs for attention is what sort of education system would back teachers and administrators into such a tight corner that they would cheat to ensure they have work in the future? It is this set of questions that ought to be debated, rather than trying to answer the question most often heard, namely, how can these miscreants be fired?

There is still another way to think about the frequent cases of cheating and the endemic violation of the standardization procedures, both of which cause serious problems in interpreting the scores of unaffected and affected students, alike. Suppose that the cheating and violations of standardization are as often acts of civil disobedience or resistance as they are of malfeasance? In reading the examples presented in Table 1, it

This document is available on the Education Policy Studies Laboratory website at:
http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc

became plausible to us that teachers and administrators are acting no different than those who have not reported all their income to the IRS, allow prayers to be said in the schools, or defy laws that deny full equality for ethnic and racial minorities or woman. In each of these disparate cases sizable numbers of people, some liberal and some conservative, decide that their government is wrong and find justification to break laws they consider unfair. We certainly should not condone civil disobedience, but we need also to look considerately at why large numbers of taxpayers cheat on their tax returns, why some religious people purposely break state and federal laws, and why some woman and minorities are impelled to protest certain actions.

It is plausible that teachers and administrators are trying to resist a system they see as corrupt and unfair, as do tax, religious and civil rights protesters across this nation. One example of this is the North Carolina principal who will not test what she calls "borderline kids," her special education children, despite the requirement to do so.[36] She says, "I couldn't. The borderline children experience enough failure and do not need to be humiliated by a test far beyond their abilities." By not testing all the children in the school the principal is cheating. But this is also an act of human kindness. And it is at the same time an act of resistance to laws made by policy makers in some other community. It is not easy to judge this principal harshly.

It is plausible that teachers and administrators are trying to resist a system they see as corrupt and unfair.

Moreover, resistance of this kind is not hard to imagine after learning that some of the nation's most honored and respected measurement and statistical experts say that the

This document is available on the Education Policy Studies Laboratory website at:
http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc

current high-stakes testing accountability system cannot work. For example, the former president of both the National Council on Measurement in Education (NCME) and of the American Educational Research Association (AERA), Robert Linn, has said the testing systems we use cannot do what the politicians want them to. He says, "Assessment systems that are useful monitors lose much of their dependability and credibility….when high stakes are attached to them. The unintended negative effects of the high-stakes accountability….often outweigh the intended positive effects."[37]

Robert Brennen (2004), who is the E. F. Lindquist Chair in Measurement and Testing, Director of the Center for Advanced Studies in Measurement and Assessment at the University of Iowa, and also a past president of the NCME said recently, "The accountability provisions of [NCLB] and its regulations are *outrageously unrealistic and poorly conceived* from a measurement perspective…. For the most part, I would argue that a more accurate title for the Act might be '*Most* Children Left Behind.'"[38] (emphasis added)

Stephen Raudenbush (2004) is a Professor of Education and Statistics at the University of Michigan, one of the leading statisticians in the nation, and also a member of the prestigious National Academy of Education (NAE). In a recent speech to the Educational Testing Service (ETS), one of the developers of tests that have high-stakes attached to them, he said: "High-stakes decisions based on school-mean proficiency are scientifically indefensible. We cannot regard differences in school mean proficiency as reflecting differences in school effectiveness."[39] Raudenbush also said "To reward schools for high mean achievement is tantamount to rewarding those schools for serving students who were doing well prior to school entry."[40]

This document is available on the Education Policy Studies Laboratory website at:
http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc

When so many prestigious scientists say the system is not working, is unfair, punishes people that should not be punished and rewards people that should not be rewarded, then the groundwork is laid for resistance, passive aggression, or civil disobedience. Some of the stories in Table 1, which follows, seem to reflect this attitude.

This document is available on the Education Policy Studies Laboratory website at:
http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc

Table 1: Instances and Allegations Cheating by School Personnel

Location of Story	Source	Headline	Story
1. New York	Daily News, Alison Gendar (November 13, 2002), p. 3.	State tests on hold till HS erases cheat label	A Boerum Hill alternative high school is barred from administering state tests until it proves its staff will not help students cheat. Helen Lehrer, the former principal at Pacific High School in Boerum Hill, was accused of erasing and changing as many as 119 answers on 14 Regents competency tests in global studies and U.S. history given last June.
2. Austin, Texas	San Antonio Express-News, Roddy Stinson (September 17, 1998), p. 3A.	TAAS cheaters meet national standard	The Austin School District manipulated test results last spring to make it appear as if several schools performed better than they did, the Texas Education Agency says. "There must be 50 ways to cheat on TAAS tests, and administrators at several Austin schools employed one of the most clever chicaneries in manipulating history. Commissioner of Education Mike Moses explained the trickery in this August 14 letter to the school district: "...student identification number changes were submitted for students tested at (the schools), which resulted in the exclusion of those students from the accountability subset of TAAS results used to determine the 1998 accountability ratings." In plainer English, administrators gave students who performed poorly on the test ID numbers that did not agree with previous numbers, knowing that the inconsistencies would cause the TEA to eliminate the students' scores from ratings calculations.
3. Potomac, Maryland	Associated Press & Local Wire (June 1, 2000) and Associated Press & Local Wire (June 6, 2000).	Principal resigns over allegations of cheating & Principal takes blame for cheating scandal on state test	First story about allegations that a principal of an elementary school ranked third in the state on the MSPAP helped students with the tests. Parents had complained that students received "inappropriate assistance intended to boost their scores" during spring, 2000 assessments. In a follow-up article, the principal took full responsibility for the allegations that led to an investigation and evidence that the principal coached students, gave them answers and extra time to complete the assessment.

This document is available on the Education Policy Studies Laboratory website at: http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc

Table 1, Continued

4. Kentucky	The Herald-Leader, Linda B. Blackford and Lee Mueller (Staff writers) (March 22, 2004).	Former principal expected to be named county superintendent	In Bell County, the school board is expected to choose as its new superintendent a former high school principal whose certification was suspended due to allegations of cheating on a statewide assessment test. According to the story, the principal in 1993 won the Milken Family Foundation Award for improvements at Bell County High School. The national award is given every year to outstanding educators, along with a $25,000 check. However, three years later the high school statewide test scores spiked so sharply that local and state officials decided to investigate. "Investigators found that among some 80 violations, Thompson had encouraged inappropriate practices, such as teachers developing tip sheets, and students had been encouraged to seek help from teachers, according to investigation and court documents." As a result, the principal was stripped of all his teaching and administrative certificates for 18 months. Although he fought the charges all the way to the Kentucky Supreme Court, he inevitably dropped the case and accepted the penalty in 2002. He is not eligible to receive his certificates back until July 2004, however, he seems to be the front running candidate for the superintendent position in spite of the "bad rap." Interestingly, the principal, Thompson, has worked for a construction company since he lost his certification. While he fought the case, he was installed as testing director for Bell County.
5. Chicago, Illinois	Chicago Sun-Times, Rosalind Rossi (May 5, 1996), p. 13.	Grade school under investigation as teachers allege test cheating	One elementary school's principal and curriculum coordinator were accused of giving teachers copies of the Iowa Test of Basic Skills and telling them to use it to help students prepare for the exam. Although the chief executive officer said he did not believe the alleged cheating was linked to stakes attached to that year's ITBS, the timing couldn't have been more coincidental as that year was the first year stakes were attached to the ITBS.
6. Maryland	Washington Post, B. Schulte (June 1, 2000).	School allegedly cheated on tests	Principal stepped down amidst charges she "was sitting in the classroom, going through test booklets and calling students up to change or elaborate on answers."
7. Wyoming	The Associated Press (July 15, 2000).	WyCAS tests lost from school at center of tampering scandal	In June of 2000, school officials said as many as 90 national TerraNova tests completed by first, second, and third graders had been tampered with, nullifying half the results. "Principal Jean Grose resigned in the wake of the finding, which resulted in scores for one class jumping from 42 percent to 87 percent over one year."
8. Georgia	The Atlanta Journal and Constitution, Michael Weiss (November 9, 1995).	Clayton school officials deny pressure to succeed	An investigation by the Professional Practices Commission, the state agency that polices teacher ethics, concluded that North Clayton Middle School Principal Gloria Duncan showed teachers advance copies of the state Curriculum-Based Assessment test and even drew up a list of vocabulary words from the exam.

This document is available on the Education Policy Studies Laboratory website at: **http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc**

Table 1, Continued

9. Arizona	The Arizona Republic, Anne Ryman (October 8, 2003).	District audits schools after test deceit	The state's eighth-largest school district is auditing all of its schools after officials alleged that a principal changed test scores so that her teachers could get incentive money.
10. Carthage, North Carolina	News and Observer, Associated Press (October 17, 2003).	Moore County SAT fund probed	Allegations that school administers were ordered by the superintendent to alter and destroy documents describing an SAT incentive program at North Moore High School. The documents showed that students who performed well were to be paid, but the money was improperly used to pay counselors. (see also table 2).
11. Atlanta, Georgia	Associated Press (May 27, 2003).	School officials suspended after cheating allegations	Two Worth County school administrators were suspended without pay following allegations that they helped three students cheat on the high school graduation test. The students were not allowed to participate in the graduation ceremony.
12. Massachusetts	The Herald Tribune, Steve Leblanc (Associated Press writer)(April 9, 2004).	Worcester school principal resigns in wake of MCAS cheating probe.	A Worcester elementary school principal submitted her letter of resignation to the superintendent of schools a year after allegations that she helped students to cheat on the MCAS exam. The principal, Irene Adamaitis, was alleged to have distributed the test to teachers days before they were administered to students. The dramatic improvement in MCAS scores at this school had prompted the investigation by the Worcester school system and the state's Department of Education. The allegations are still under investigation.
13. Boston, Massachusetts	National Post, Mary Vallis (staff) (June 24, 2004).	The "principal" told us to: Critics blame Boston cheating scandal on pressure to score high in standardized tests	Students told investigators that Antoinette Brady urged them to cheat on a high-stakes standardized test. They had written to their teacher alleging that the principal had encouraged them to change and add to their answers after the exam was officially over--the opposite of the instructions their teacher had given them. One student had written, "when we were done Ms. Brady told us to go back to Section 1 and 2. So everybody was told by the PRINCIBLE." "We did it. She look into everybodys eyes and said did you go back to the sections and check it over." One student reported that "Ms. Brady said that are school had low performance and we need it to be higher" These allegations emerged Friday, May 21 when one fourth-grade class wrote for the MCAS exam. The principal had asked the teacher to leave the room.

This document is available on the Education Policy Studies Laboratory website at: **http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc**

Table 1, Continued

14. New York	The New York Times, Randal C. Archibold (December 8, 1999).	Teachers tell how cheating worked	Story on one of the biggest cheating scandals in New York City. A year after he arrived at Community Elementary School 90, Jon Nichols, a mathematics teacher, was approached by the principal, Richard Wallin, with what seemed at first an unusual pep talk. “I was taken into the office at the beginning of the school year and told that the students were expected to do well -- no matter what it takes,” Nichols recalled Tuesday at the school in the Concourse section of the Bronx. “At the time, I didn't know what that meant exactly.” He learned soon enough. It meant cheating, and Nichols and other teachers were provided with detailed instruction, down to palm-sized crib notes to check against students' answer sheets as they took city and state examinations, according to accounts from Nichols, other teachers and Edward F. Stancik, the special investigator for New York City schools. Stancik released a report Tuesday asserting widespread cheating on city and state exams over the last four years. The report identified 32 schools, but cited Community Elementary Schools 88 and 90 in District 9 in the Bronx and Public School 234 in District 12 in the Bronx as particularly egregious cases. The investigation began with Community School 90 after a teacher, Stacey Moskowitz, contacted Stancik's office and a television news reporter about the practices. Nichols and others then came forward. Nichols, 33, who investigators confirmed had participated in the case, described how an administrator approached him before city reading and math exams in April 1994. "Keep this handy, he said, and he gave me a piece of paper," Nichols recalled. "It was a 2-by-3 sheet with a list of numbers and letters on it. It was the answers to the test. Some of the numbers had asterisks next to them," Nichols continued. "These were the hard questions, and I was told not to help with those. The kids were expected to get them wrong."
15. Houston, Texas	Associated Press, Pauline Arrillaga, Associated Press Writer, (April 13, 1999).	Tampering allegations raise new questions about the role of TAAS	Story arguing that the pressures associated with testing are so great, it tempts teachers and administrators to do whatever it takes to look good. In the largest school district in Houston, administrators launched an investigation into the high number of erasures turning up on TAAS tests, with answers being changed from wrong to right. This probe (along with one in Austin) led the Texas Education Association to initiate its own investigation of the erasure problem—asking 11 of the state’s 1,042 school districts to investigate 33 campuses that had above-average erasures in 1996-1998. Several Houston area districts found cheating or improper administration of TAAS at 11 campuses. “The infractions ranged from removing partially erased answers to using keys to ‘systematically check’ and change answers from wrong to right.” One fifth-grade classroom had 185 erasures attributable to only 14 students, including 132 that were changed from wrong to right. A principal and a teacher in another district resigned, and Houston demanded the resignations of a principal and three teachers.

This document is available on the Education Policy Studies Laboratory website at: **http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc**

Table 1, Continued

16. Milwaukee, Wisconsin	Milwaukee Journal Sentinel, Sarah Carr (Staff Writer) (January 30, 2004).	MPS looks into claims staff helped on tests	One elementary school principal and two teachers were accused of giving students answers to questions on the statewide assessment (the Wisconsin Knowledge and Concepts Examination and TerraNova), both taken in November 2003. The article notes that the stakes attached to the Wisconsin assessment are possible motives for the alleged aid. The allegations are that the teachers "had students memorize the sequence of answers" on a multiple-choice section of the test. Allegations involved students in third, fourth and fifth grades. The school in which the allegation occurred has low test scores. In 2000, only 23 percent of third graders were in the top two brackets, and 32 percent drew the lowest ranking ("minimal"). By comparison, four percent of students statewide were ranked as "minimal."
17. Maryland and Virginia	Washington Post, L. Seymour; W. Branigin (June 2, 2001), p. B2.	Theft forces school to give state test again	Seven teachers and administrators were removed from their posts when some students had seen test questions before they took their achievement test of basic skills. And the year before, in the Fairfax, Virginia, two teachers resigned after being accused of reviewing test questions with students prior to the test.
18. Montgomery, Alabama	Associated Press (April 21, 1999).	Superintendent: Five schools investigated for cheating	Story about how the state was investigating allegations of cheating by a superintendent and some teachers in reaction to higher than expected SAT scores. According to the article, "Education Department spokesman Tony Harris said the state always has a few investigations into the achievement test, which is used as a measurement of academic performance by public schools and school systems. 'It's not uncommon, that almost without exception, you'll have an isolated incident or two of irregularities,' Harris said. The test is 'a big indicator for alert status,' he said. Schools are put on 'academic alert' when students average low scores on the achievement test. Alert schools are at risk for state takeover if they don't improve." It is believed that making the test an accountability factor could lead to more "irregularities."
19. Greensboro, North Carolina	Bruce Buchanan, News & Record (February 21, 2003), p. A1.	Teachers may earn bonuses despite scandal	Story about teachers at Andrews High School who will still receive their performance bonuses in spite of allegations of cheating that led to the resignation of a few teachers. Three teachers and a central office testing coordinator resigned after allegations that they were accused of sharing physical science and U.S. history end-of-course test questions with students before giving them the test in January. Another teacher received a five-day suspension for distributing the test—a practice he didn't know went against district/state policy.

This document is available on the Education Policy Studies Laboratory website at: **http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc**

Table 1, Continued

20. Louisiana and Connecticut	Times-Picayune, Jeffrey Meitrodt (September 17, 1997).	LA Voids scores in 19 schools: Erasures blamed	A suspiciously high number of erasures on standardized-test answer sheets prompted the State Department of Education to throw out scores from 10 Orleans Parish public schools. Regulators were keeping an eye on erasures because they can help identify classrooms where teachers are improperly helping students. Infractions can range from a teacher coaching a student during the test, to a principal erasing students' wrong answers and filling out the right ones. According to New Orleans Schools Superintendent Morris Holmes, "We are investigating as if there were improprieties," Holmes said. "We have a whole lot of erasures that statistically could not just happen by chance. So we are trying to determine how that happened." One of the biggest education scandals in the country involves a nationally recognized public school in Fairfield, Conn., where the principal was suspended after state officials noticed a suspiciously high number of erasures on students' standardized tests.In the Connecticut case, 89 percent of the erasures resulted in answers being changed from wrong to right.
21. Seattle, Washington	Seattle Times Linda Shaw and Tan Vinh (Staff Writers) (November 26, 2003).	Three Franklin counselors disciplined over grade changes	Three counselors inappropriately changed student grades in the '02-'03 school year at one high school. The district's policy was that students must get at least a 2.0 or C average in order to graduate. However, an investigation into the allegations found that the counselors were not alone in changing grades. The school as a whole had adopted grading policies that conflicted with district-level rules under a previous principal. Therefore, counselors and teachers who changed grades to help students graduate were in violation of district-level policy, but within the approved school-level one.
22. New York	New York Times, Abby Goodnough (December 10, 1999).	Union questions investigator's finding on teacher cheating	A Manhattan District Attorney accused 43 teachers and two principals in 32 schools for participating in some form of cheating. In this accusation, teachers were accused of changing kids' answers on standardized tests from wrong to right—in part because teacher pay and monies allotted to the schools were contingent on students' test performance.
23. New York	Daily News, Paul H. B. Shin (December 11, 1999).	Crew lowers boom: Axes 9 teachers, aides who helped kids cheat	Schools chancellor Rudy Crew fired nine educators and put another 11 under the ax for helping students cheat on high-stakes standardized test. Those who were fired included three paraprofessionals, or teachers' aides, and six provisional teachers who did not have tenure. The board's action came three days after special schools investigator Ed Stancik released a bombshell report accusing 52 teachers, principals and other staff at 32 elementary and middle schools with aiding and even encouraging cheating on city and state standardized tests. Stancik described methods of cheating – from giving students the answers outright to suggesting they redo a problem. In one case he cited, the teacher simply added a paragraph – in her own handwriting – to a fourth-grader's essay for the state English exam. Among the 52 named by Stancik, seven had resigned or been dismissed before the report was released, according to board attorneys.

This document is available on the Education Policy Studies Laboratory website at: **http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc**

Table 1, Continued

24. Delaware	Daily Times, Editorial (June 23, 2004).	High stakes, big problems: Lessons in Seaford's errors	Editorial writer argues that the pressures of high stakes testing are creating an environment making it more tempting for teachers to cheat. Because of the severity of the penalties for low scores or for failing to increase them, jobs can hinge on the annual scores. So preparing students for the tests is increasingly a priority for both teachers and administrators. And that is what led to a problem in the Seaford School District this spring. It is common practice to hand out sample questions from previous tests to help prepare students to score well on the exams. But in Seaford, copies of the current exam found their way into the hands of three tenth-grade English teachers, one of whom reported the similarity after the fact, when she realized the samples and the actual test were identical. The samples had been distributed by the principal, who at first was accused of deliberately providing early copies of the current test. Later, it was determined the incident was a mistake rather than an attempt to cheat. But it could not be ignored, especially since the principal, Michael Smith, had been serving as the testing coordinator for his school. The school board decided the scores of students in the two tenth-grade classes whose teachers had actually used the samples would not count in rankings or averages for the state. But on an individual level, any student who qualified for a scholarship would still receive the assistance, since the error was not the students' fault. To make it fair, for each scholarship awarded to a student in one of those classes on the basis of that test score, another scholarship would be awarded to the next eligible contender on the list. Smith will foot the bill for those extra awards, up to $5,000 worth. And Smith will be responsible (up to $5,000) for any legal fees associated with his mistake. He also will no longer serve as state testing coordinator. The Seaford incident is one of three such incidents deemed "serious" by the Delaware State Board of Education this year alone. If there is a moral to this story, it is that high-stakes testing is vulnerable to human error as well as the overwhelming urge to meet the standards and score well at almost any cost. Even without cheating or errors in judgment, some say that education itself suffers because preparing for a single test can dominate the curriculum for an entire school year, cheating students of opportunities to benefit from creative teaching methods or opportunities to learn beyond the scope of the testing.
25. South Bend, Indiana	South Bend Tribune, Editorial (June 18, 2000), p. B10.	Teacher cheating is sign of sick system	Editorial lamenting the pressures faced by teachers to increase student achievement measured by standardized tests. It is no surprise some teachers cheat given that financial incentives are tied to performance.
26. Tennessee	The Commercial Appeal, Editorial (September 24, 2003), p. B4.	High-stakes testing turns the screw	Article about the effects of the pressures of high-stakes testing on teachers and students. In addition to forcing teachers to narrow their curriculum, a survey of teachers submitted to the newspaper reveals instances of cheating/fudging data for schools, students, and teachers to look better on tests. “Those who had witnessed or took part in such cheating – about nine percent of teachers who responded to the newspaper's questionnaire – cited instances in which weak students were held out of class on test days to raise average scores, students were assisted during testing and answers were changed. The cheating, although small in scale, is a predictable outcome of new sanctions that could affect educators' jobs and reputations as well as local school district autonomy and funding.”

This document is available on the Education Policy Studies Laboratory website at: **http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc**

Table 1, Continued

27. New York and National perspective	New York Post, Carl Campanile (January 26, 2004).	Teachers cheat	Story about teacher "scrubbing" – the process of "tweaking" student scores on tests (namely, the statewide Regent's test). The article says that scrubbing is most often done on the English and history exams because they have essay questions that are subjectively scored. In this process, a student's exam is reviewed by at least two different teachers. The scores these two teachers assign are averaged (required under state rules). Then, these papers are sent back to respective departments, where some teachers, often under the guidance of principals, set aside tests that are just a few points shy of passing. Then, teachers review exam responses and "find" extra points to assign to students. This process is heavily debated. Some say it is not cheating. According to one teacher: "I'm sorry if it's shocking for layman to hear. Scrubbing is something we do to help the kids get their asses out of high school." Others argue it is blatant cheating. And, there are plenty of incentives to do so with merit pay bonuses and diplomas on the line. According to one high school staffer, "The students of the school benefit because they pass. The school benefits because the pass rate is up…but it is grade inflation. It's cheating. You're falsifying exams. It's totally corrupt." The story's writer discusses other instances of cheating found in an Internet search: From the (New York) Daily News: "Teachers at a South Bronx elementary school say their principal helped hundreds of students cheat on citywide math and reading tests." (A witness) said he saw school staff open the sealed test, complete it and transfer the correct answers to yellow sheets, which were given to the teachers before the test. "During the test, the teachers made sure students were entering the correct answers." Reports taken from the (New Orleans) Times-Picayune: "One New Orleans teacher has resigned, a second is under investigation ... in the wake of questions about erasures on standardized tests their students took in the spring … Statewide, at least five teachers linked to suspected cheating have resigned or are being fired, and two others have been reprimanded." Report from the Montgomery (Alabama) Advertiser: At Sumter County High School, "at least 100 students" obtained Stanford Achievement Test booklets "that may have been the source of cheat sheets." As a result, "students who had ranked in the 38th percentile in mathematics scored in the 97th percentile." A whistle-blowing teacher "was fired by the school board." From the Seattle Times: "There were reports of teachers and other school staff members copying secure test materials to share with students, reading passages on a reading section of the test to students and correcting students' spelling on the writing section."
28. Nevada	Las Vegas Sun, Associated Press (January 18, 2004).	Student testing irregularities increase in Nevada	Nine cases of student cheating on standardized tests were reported statewide last year. Four cases were alleged security violations by teachers. The education department reported 79 testing irregularities statewide in 2002-2003, up from 50 the previous year. Irregularities include—power outages and missing test booklets to student misconduct and improper test administration. Six of the nine instances were on the high school proficiency exam – "likely because the stakes are higher," according to the states assistant director of accountability.

This document is available on the Education Policy Studies Laboratory website at: **http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc**

Table 1, Continued

29. California	Los Angeles Times, Erika Hayasaki (May 21, 2004).	One poor test result: Cheating teachers	Since the statewide testing program began five years ago, more than 200 California teachers have been investigated for allegedly helping students on state exams, and at least 75 of those cases have been provided. Cheating behavior included: Whispering answers into a student's ear during an exam; photocopying test booklets so students would know vocabulary words in advance; and erasing answers marked with wrong answers and changing them to correct ones. Teachers receive enormous pressure from principals to work on raising scores – not just for bragging rights, but because federal funding can be withheld. Some incidents in the last five years include: • In the San Joaquin Valley's Merced County, a third-grade Planada School District teacher gave hints to answers and left a poster on a wall that also provided clues; • In the Inland Empire, a Rialto Unified School District third-grade teacher admitting telling students, "You missed a few answers; you need to go back and find the ones you missed." A student reported that the teacher looked over pupils' shoulders and told them how many questions were wrong; • Near the Mexican border, in the El Centro Elementary School District, a principal asked a student why he had erased so many answers. The student responded that the teacher had told him to "fix them." • In El Monte, a Mountain View School District eighth-grade teacher admitted using the board to demonstrate a math problem and saying, "This is a silly answer. If you marked this one, erase it and pick another." Records stated that the teacher, "said she was very sorry and wept during the interview." • In the Ontario-Montclair School District, a student told investigators that a teacher read 10 math answers. One student said he handed his test booklet to that teacher and then went back to change five answers after the teacher said, "Why don't you try again?" • Near Salinas, a Hollister School District teacher admitted changing about 15 answers. Beverly Tucker, California Teachers Association Chief Counsel for 16 years said the number of teachers her office defended against allegations of cheating had risen. She could recall one or two cases stemming from the decade before the current testing began. Since 1999, she estimated the union has defended more than 100.

This document is available on the Education Policy Studies Laboratory website at: **http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc**

Table 1, Continued

30. California and Maryland	Fresno Bee, Angela Valdivia Rush (July 26, 2000).	Teachers disciplined in cheating	Three Central Unified teachers and at least three Fresno Unified teachers have been disciplined for helping students cheat on the Stanford 9, a statewide, high-stakes test used to measure students' achievement. A Central Unified teacher, according to district officials, allowed an entire class to use multiplication tables while taking the math portion of the SAT 9 – a violation of the test's guidelines. The teacher was exposed by a student. After an investigation, the district discovered students in three classrooms were allowed to use multiplication charts as their teachers stood by. In Fresno Unified, at least three teachers were disciplined after it was discovered they made copies of the SAT 9 test and used them as study guides, according to a high-ranking Fresno Teachers Association official who requested anonymity. One of the more highly publicized instances of cheating was at a top-ranked school in Maryland. In June, the school's principal resigned and a teacher was placed on administrative leave after being accused of rigging a statewide achievement test. The pressures are no different in California. Under the state's new accountability plan, if SAT 9 scores improve, schools could qualify for extra money for programs and teachers could receive bonuses of up to $25,000 for personal use. If scores take a dip or remain stagnant, a school could face sanctions that include removal of the principal or takeover by the state.
31. Michigan and National perspective	The New York Times, Jodi Wilgoren (June 9, 2001).	Possible cheating scandal is investigated in Michigan	Michigan school officials investigate irregularities on the state's standardized tests in what appears to be the "largest of a recent spate of cheating scandals across the country as the use of high-stakes tests has risen. At least 71 schools – most in Detroit but scattered among 22 districts in the state – have been alerted to possible cheating after test graders found remarkable similarities among students' written responses on science, social studies and writing exams, which determine how millions of dollars in scholarships and school awards will be distributed." In some cases, several students in the same classrooms submitted paragraphs that were practically identical to open-ended prompts, raising questions of whether they inappropriately worked together or were even fed answers by adults. "The accusations join a growing list of investigations in at least half a dozen states in what experts see as an inevitable outgrowth of the expanding use of standardized tests to determine promotion, scholarships, teacher pay and school financing. Even more such testing is promised since Congress passed President Bush's education plan, which requires annual exams in grades three through eight." Several teachers have been suspended and hundreds of students forced to retake tests in Virginia this month after exams were circulated in advance. In Maryland, Montgomery County is spending more than $400,000 to replace a test that a teacher photocopied and distributed, while a principal at a different school resigned after being accused of giving students extra time and coaching them to change answers. A dozen school districts in Texas are looking into patterns of excessive erasures and corrective editing. In 1999 in New York City, dozens of teachers and two principals were accused of cheating on reading and math exams in 32 schools, though later investigation suggested reports were overblown.

This document is available on the Education Policy Studies Laboratory website at: **http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc**

Table 1, Continued

32. Utah	The Desert News, Jennifer Toomer-Cook (Staff Writer) (August 10, 2002), p. A1.	Teachers cheating in record numbers	Utah's reported incidents of teacher cheating are scattered throughout the state, as reported by the state's testing director. Accusations that have been recorded include: A teacher who was accused of changing students' answer sheets; another teacher who passed tests to colleagues and talked about how they could teach to the test questions; one teacher who marked certain questions on the core curriculum test and told students to carefully think about them before answering; another who changed students' scores on college entrance tests, allegedly to notch up the class' report to the school district. "I think they're concerned with how public all of this information is, and wanting to look good," Lawrence said. "Some of it, however, is teachers not fully understanding what is and is not acceptable … and in some cases, it is a matter of teachers being truly naïve." That could be the case with one teacher who allowed students to work on the Stanford Achievement Test during recess or lunch breaks. The SAT is a timed test.
33. Chicago, Illinois	Chicago Sun-Times, Rosalind Rossi; Kate N. Grossman (October 3, 2002), p. 7. Chicago Sun-Times, Rosalind Rossi & Annie Sweeney (October 2, 2002), p. 1.	Pressure blamed in test scandal	It was reported in both articles that in 11 of 12 Chicago Public School classrooms under investigation for cheating, the "stakes were so high that kids who flunked their tests would have been forced to attend summer school." All teachers under investigation taught in "critical" grades (third and eighth), where students could be held back if they scored poorly. Teachers were accused of giving tips during the ITBS in May, erasing incorrect answers, and filling in answers that had been left blank.
34. Buffalo, New York	Buffalo News, Darryl Campagna & Charity Vogel (August 29, 1999), p. 1A.	Teachers under pressure: New goals may tempt more faculty to cheat	In Buffalo, a teacher resigns after being accused of giving his students advanced copies of questions from a district wide test. Six weeks later, an Amherst teacher resigns after an investigation found that he tampered with 48 Regents test scores sent to the state. Article notes that the pressures are greatest on younger, untenured teachers.

This document is available on the Education Policy Studies Laboratory website at: **http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc**

Table 1, Continued

35. New York	Stancik, E. F., & Brenner, R. M. (1999, December). Special report of the Special Commissioner of Investigation for the New York City School District.	Cheating the children: Educator misconduct on standardized tests	Special report released by the Stancik Commissions outlining allegations and indictments of teachers who had cheated. The report outlines the various ways teachers were found to cheat. For example, one teacher instructed students in a fourth-grade class to write their answers on a piece of loose-leaf paper which they then submitted to be "corrected." In another example, one third-grade teacher had students place their answers on a "green" bubble sheet first. Then, the teacher roamed around the room, stopping at desks and pointing to or stating the number of those that were wrong. The children then "checked these selections and picked new ones." Then, they transferred the answers to the official "pink" bubble sheet. In another example, one third-grade teacher pointed to wrong answers on the 1998 math test being taken by third graders and told students to "try to get the answers right before putting them on the grid paper." In other cases, teachers outright gave students the answers. For example, during the eighth grade math exam in 1996, one teacher gave students answers. According to one of the students, he explained "how to do the equation," and he "helped almost everybody."
36. New York	Associated Press (October 29, 2003).	Teachers cheating blamed on high-stakes testing	From 1999-2001, New York education officials found "21 proven cases of teacher cheating from Buffalo to Long Island." Teachers have been accused of "reading off answers during a test, sending students back to correct wrong answers," and for photocopying secure tests for use in classes.
37. New York and Nationwide	Associated Press. October 28 2003. http://www.azcentral.com/news/articles/1028CheatingTeachers28-ON.html	AP Exclusive: Records show teachers cheating on tests to boost scores	Story about how the pressures of high-stakes testing had led to "high-level cheating—by teachers trying to gain an advantage for their students." Article comments on how from spring 1999 through spring of 2002, New York education officials found 21 proven cases of Teacher cheating from Buffalo to Long Island. Some teachers said that it is even more extensive than that. Other similar stories found: in Austin, TX, grand jury found 18 school officials guilty of altering tests; in Columbus, Ohio, school praised by President Clinton had adult tutors guiding students' pencils and calculating math problems during the test; etc. Cheating is so endemic that it has invalidated the school report cards relied on by parents and the state to evaluate schools.
38. New York	Daily News, Paul H.B. Shin (May 3, 2000), p. 6.	New teacher test mess Stancik: Rotten apples pushed kids to cheat	Article alleges that some New York City teachers have lured students into cheating on tests. Stancik, an independent investigator of the New York City teachers during a cheating scandal, alleges that some teachers were pressuring students who refused to accept their help. Said Stancik: "It sends a terrible message in terms of [students'] moral compass when they're seeing not only teachers who will cheat, but teachers who are trying to lure them or even force them into cheating."

This document is available on the Education Policy Studies Laboratory website at: **http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc**

Table 1, Continued

39. New York	Michael Gormley, Associated Press, (October 28, 2003) & (October 26, 2003).	AP Exclusive: Records show teachers cheating on tests	Article reports on a series of teacher cheating cases. The records include evidence that: • Teachers ignored cases of cheating by others. • Colleagues protected cheaters in statements that were later recanted. • Some teachers shared the secure information with others and helped protect offenders with union involvement instead of notifying administrators. • A teacher at Putnam County's Mahopac High School reported all his students passed the December 2000 Regents chemistry exam. An investigation found 62 of 63 of the exams were scored higher than deserved and 16 students failed. • A teacher at the Washington-Saratoga-Warren-Hamilton-Essex BOCES coached students through a Regents Competency Test in global studies test in January 1999 and even wrote in correct answers. The same teacher was accused of cheating in the January 1999 math Regents Competency Test by telling a student what questions were wrong. "I've done it before and this isn't the first time people have accused me of cheating," the teacher said. • A teacher at the Huth Road Elementary School of Grand Island, while visiting his former school in Buffalo, looked at the fourth-grade English test while forms were being counted. The teacher shared the information with colleagues, including one who wrote a 26-page report of the ordeal that followed. She said she agonized with colleagues for days over whether to report the incident, fearing her involvement might hurt her chances to become an administrator and damage the district's integrity if parents found out. Ultimately she passed an anonymous note to the principal. "Everyone else seemed comfortable burying their heads in the sand," she wrote. • A Furnace Woods Elementary School teacher in Montrose coached students in the May 2002 fourth-grade math standardized test, including telling them when they made an error. The teacher said she "does this all the time during tests and saw nothing wrong with it." "Teachers are under a lot of pressure to get good grades," she told administrators.

This document is available on the Education Policy Studies Laboratory website at: **http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc**

Table 1, Continued

40. South Carolina	The Herald, Staff (July 26, 2002)	Leaked info from PACT costs state money, time	Allegations surface that questions from the state's standardized test have been leaked to the public. According to the report, each time a question is leaked, it costs taxpayers $2,500 and more than two years to fix. In the past year alone, "10 security breaches have become known with PACT in Oconee, Greenville, Spartanburg District 5, Anderson District 5, Laurens District 55 and Darlington. SLED is investigating some of the cases, and solicitors are reviewing some of them." Barbara Parrish, assistant superintendent for elementary programs and services and test coordinator for the Clover school district, said she believes that because accountability measures have become an even greater focus statewide, teachers may feel more pressure. "It is unfortunate when these things begin to happen and it is a scary thing when security breaches have to be reported to SLED," Parrish said. The case in Darlington involved a husband and wife in schools in Darlington and Florence School District 1, respectively. According to Hodge, the wife asked her husband over e-mail what the PACT writing prompt was. He sent it back and she shared it with her students. A student turned her in, Hodge said. This fall, the saga to replace the Darlington items will begin.
41. Providence, Rhode Island	Providence Journal-Bulletin, Elizabeth Rau (March 13, 1999), p. 3A	State may report on test probe next week	On Wednesday, March 10, 1999, it was announced that an investigation was being launched into allegations of teacher cheating and a breach of professional conduct. This story is a follow-up story saying that 36 school districts had filed reports detailing their involvement (or lack thereof) in the teaching scandal.
42. Pennsylvania	Pittsburgh Post-Gazette, Eleanor Chute (May 25, 2003), p. 1.	Pupils prompt cheating probe; Popular 4th grade teacher suspended over math testing	One 4th grade student reported to his mother that he is sick of taking his math tests over and over again. When his mother asked what he meant, he reported, "the teacher checks the tests, and then gives them back to the class with sticky notes on them. The notes give clues about answers that are wrong. And, she also talks to them about the notes."
43. Pennsylvania	Associated Press, Martha Raffaele, (September 18, 2003)	Lawsuit alleges teacher was fired for reporting test cheating	A former teacher at a Philadelphia charter school filed a lawsuit Wednesday that claims she was wrongfully fired three days after she told city school officials that administrators helped students cheat on state assessment tests. According to her complaint, she contacted the district office on March 25 after she saw an administrator tell third-grade students who were taking the test to stop and check their work. The administrator, who was not identified in the lawsuit, also allegedly looked through one student's test book.
44. Worcester, Massachusetts	Telegram & Gazette, Clive McFarlane (Staff) (December 15, 2003), p. B1.	MCAS allegations shock parents; Chandler School high scores probed	Allegations that teachers at Chandler Elementary School may have inappropriately helped students on the MCAS. Parents believed the high scores on the exam were the result of hard work and were very surprised to hear that kids may have been helped. Allegations were that teachers coached students as they took the exam. But, while cheating might seem anathema to the teaching profession, it is not far-fetched, and parents and the community will have to face that reality, according to S. Paul Reville, Executive Director of the Center for Education Research and Policy at The Massachusetts Institute for a New Commonwealth. "The fact that there are incidents or allegations in response to high-stakes testing indicates that real performance pressure now exists in public education," he said.

This document is available on the Education Policy Studies Laboratory website at: **http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc**

Table 1, Continued

45. Chicago, Illinois	Chicago Sun-Times, G. Wisby (January 23, 2003), p. 11.	Aide fired in test cheating scandal	An aide was accused of helping students to cheat on the Iowa Test of Basic Skills exam. It was alleged that the aide pointed students to correct answers and urged students to re-read questions they answered incorrectly.
46. Birmingham, Alabama	Birmingham News, Charles J. Dean, Julie Ross Cannon and Jeff Hansen News (Staff writers)(June 24, 2000).	Officials to probe Stanford scores	Allegations that teachers may have been engaged in cheating because students' SAT scores were higher than was expected. According to the article, "State School Superintendent of Education Richardson said Friday that dramatic improvements in achievement test scores in some schools suggest that cheating may have occurred." Citing scores in schools across the state, the article quotes several administrators who are skeptical that these were real gains. However, some educators argued that the gains were real and the product of hard working teachers and students.
47. Nevada	The Reno Gazette-Journal Online, Associated Press (January 30, 2004).	Two Clark County teachers investigated for leaking math exam	Two high school teachers are under investigation for improperly distributing the contents of the statewide math proficiency exam to students. One teacher allegedly copied questions from the test by hand and created a worksheet that was distributed to at least two students. Another teacher allegedly photocopied and cut out portions of the test for his students to use as a practice sheet. Both incidents occurred last April in high schools.
48.Nevada	Associated Press, May 9 2004.	Nevada teacher loses his license for helping students cheat	A former Desert Pines High School teacher had his teaching license revoked after he helped students cheat on the Nevada High School Proficiency Exam. Ronald J. Williamson is the first Nevada teacher to suffer the penalty for violating high-stakes test security procedures. An investigation found that a "handwritten math study guide that he gave to at least two students contained 51 actual questions from the exam." In Nevada, all high school seniors must pass this proficiency test to receive a diploma. The study guide was discovered in an abandoned student backpack in a classroom.
49. Illinois	The State Journal-Register, Editorial, (August 14, 2003), p. 4.	Test tampering cheats district	Writer talks about the immorality of adult cheating that helps students to do better on tests. Two administrators were "alleged to have photocopied the Prairie State Achievement Exam and given it to teachers at the high school in White Hall with instructions to use the copies to prepare students for the important state testing." Teachers and administrators are feeling great pressure these days to make sure their students achieve academically. It's no longer just about looking bad. Today a low-achieving school can face serious sanctions. In fact, the new federal "No Child Left Behind" law allows parents to yank their children from a school if it repeatedly fails to show enough academic improvement.

This document is available on the Education Policy Studies Laboratory website at: **http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc**

Table 1, Continued

50. Missouri	Associated Press (July 9, 2002).	Teacher fired after admitting she peeked at MAP test	A Mid-Buchanan fourth-grade teacher has been fired after admitting she looked through the state's standardized test before it was administered to students. The teacher had been on paid administrative leave since April 23 when allegations surfaced that she helped her students cheat. The teacher admitted that she had read through copies of the Missouri Assessment Program test that were being stored in the principal's office. She told the board she was "unaware of a school policy that prevented teachers from having access to the exams before they are distributed."
51. California	San Mateo County Times, Staff (June 15, 2002).	Teachers helped pupils cheat at East Palo Alto public school	Teachers helped students cheat on standardized tests at a public school that has won national recognition for educating poor and minority students. Eleven former students at Costano School in East Palo Alto told the newspaper that teachers either gave pupils answers or strongly hinted at them. According to one former teacher, the message at the school was, "you do whatever you need to do to get those scores up!"
52. Indiana	South Bend Tribune, Dee-Ann Durbin & Ashley Lowery (Staff Writers), (February 3, 2002), p. A1.	Cheating claims alter MEAP plans	Marcellus Community schools ended up on a list of schools under scrutiny for suspected cheating on the MEAP test. The fifth-grade portion of the social studies MEAP drew suspicion because students reportedly used similar phrases or examples in two answers. "Marcellus Middle School was among 71 schools in 22 districts accused of possible cheating on the MEAP tests. As evidence, the Department of Treasury displayed essays whose answers matched nearly word-for-word."
53. Kentucky	The Associated Press (May 31, 2003).	School, state investigation allegations of classroom cheating in Whitley County	An employee of Whitley County schools has been suspended with pay over allegations of possible cheating on a state-mandated test. "The allegations, if substantiated, could cause the state to punish the school by deeming it ineligible for cash awards through the state rating system." The article does not outline the specifics of the allegations, only that preliminary investigation has revealed some improprieties in the matter in which the test was given in one elementary school. The test, the Commonwealth Accountability Testing System, is used to rate schools. "If this allegation is found to be true, the employee's education license could be suspended by the State Education Professional Standards Board. The state could also punish the school, including lowering the scores of the students who were taking the test when the infractions occurred … that could prevent the school from meeting its test goal and prevent it from qualifying for cash rewards under the state's testing and school-ranking system."
54. New Orleans	The New Orleans Channel.com, (posted April 28, 2004, http://www.theneworleanschannel.com/).	Teachers accused of helping students cheat on high-stakes test	Accusations of cheating in St. Charles parish have led to disciplinary action against three teachers. The teachers were accused of giving students answers to or providing them with the questions on the state's Iowa and other standardized tests. As a result, one teacher has resigned, and the two others were suspended. "The state Board of Education said one teacher at Norco Elementary gave out answers to the test orally as fifth-grade students took it. Another teacher at Ethel Schoeffer Elementary in Destrehan is accused of looking ahead on the test and providing sixth-grade students with questions. A third teacher at Hahnville High School is accused of copying and disseminating the questions on the high school Graduate Exit Exam." According to the article, "High-stakes test scores in St. Charles parish are traditionally among the highest in the state."

This document is available on the Education Policy Studies Laboratory website at: **http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc**

Table 1, Continued

55. Kansas	The Kansas City Star, Noemi Herrera (April 17, 2004).	Resignation follows test irregularities	A seventh-grade teacher resigned from the Olath School District after it was determined that questions on practice tests too closely resembled questions on this year's Kansas Math Assessment Test. "The tainted practice tests meant that about 1,200 seventh graders throughout the district's seven junior high schools had to retake the test." The teacher had been in the district for about 23 years. According to the president of the Olath National Educational Association, Carolyn Goode, "It would be difficult for a teacher to unintentionally design a practice test so similar to the actual state assessment that it would compromise the validity of student scores."
56. South Carolina	The State, Gina Smith (Staff Writer) (June 23, 2004).	Teacher sentences in 1st PACT cheating case: Columbia woman pleads guilty; chooses $800 fine	A Columbia woman was the first to be convicted of helping students to cheat on the South Carolina high-stakes test, Palmetto Achievement Challenge Test (PACT). This is believed to be the first such conviction since 1989. The teacher, Deborah Primus, 32, a middle school math teacher, was accused of showing and teaching actual test questions to her middle school students at the Lower Richland School. Because the state Department of Education recycles questions, it is illegal for teachers to have access to them or use them in the classroom. When the students were questioned, they said they had been taught some of the same questions in Primus' class. SLED, the enforcement agency for PACT security, also reported it found 44 PACT questions copied from PACT tests in 1999, 2000, 2001 and 2002 in Primus' classroom. "They were the exact same word-for-word, letter-for-letter, space-for-space ... questions," said Heather Weiss, Richland County assistant solicitor. Some of the graphics were also the same.
57. Arizona	Associated Press, Phoenix (July 15, 2004). Available online: www.kvoa.com	Education department investigates allegation of cheating on AIMS	Arizona Department of Education is investigating at least a dozen allegations of cheating on standardized tests. Nine Arizona school districts have nullified parts of their spring standardized test scores after allegations that teachers either read sections of the test to students or gave students extra time to finish. The state is also trying to resolve the scores of students in Phoenix's Creighton Elementary, Gilbert Unified and Yuma Elementary districts. Teachers there are accused of giving students up to three days to write essays for the AIMS test. Directions on the state test ask that students write the essay in one sitting. Other allegations include a teacher accused of helping students change wrong answers and three teachers who shared a list of vocabulary words used in the Stanford 9 test. Some school district officials say the violations are misunderstandings or lapses in judgment and not intentional cheating.
58. Arizona	Arizona Daily Star, Sarah Garrecht Gassen (July 13, 2004).	Inquiry: Did ex-teacher help kids cheat?	Story about an inquiry into whether a former Tucson Unified School District teacher told students to change wrong answers on the Stanford 9 standardized test from spring 2004.

Table 1, Continued

59. Florida	Tampa Tribune, Ellen Gedalius (June 21, 2004).	FCAT cheating penalty mild, some say	The article highlights how teachers who cheat in Florida receive different consequences depending on where they reside. The article writer notes, that recently, the three Tampa teachers who cheated on the FCAT would have been fired if they taught in Orlando. Instead, the Hillsborough County school district suspended them for two days without pay. According to the article, the state Department of Education lets districts determine how to punish teachers who help students on the high-stakes exams. Statewide, punishments range from a letter of reprimand to firing. A Hillsborough investigation released last week revealed Cannella Elementary teacher Lynette Carter gave clues and answers to students taking this year's Florida Comprehensive Assessment Test. At Gorrie Elementary, teacher Heidi Sweet put marks next to incorrect answers and teacher Jacqueline Cross tapped her finger next to incorrect answers, giving students a chance to change their responses. The students' test scores were invalidated, which could have put them at risk of being retained in third grade. This year a Leto High student who took home the FCAT was suspended and charged with three felonies, including theft and burglary. Orange County school district officials say suspending teachers isn't strong enough punishment. The Broward County school district takes a case-by-case approach. A teacher might be suspended for telling children to check their answers but stopping short of telling them which to choose. Or a teacher could be fired for collecting the students' tests and changing the wrong answers to the right ones. Broward recorded six accusations of teachers cheating on the FCAT during the 2002-03 school year and one incident in 2003-04.
60. Florida	St. Petersburg Times, Melanie Ave (July 19, 2004).	Teachers as law-breakers	Two teachers at a Tampa elementary school were caught helping third graders cheat on achievement tests.
61. Greensboro, North Carolina	Associated Press (March 18, 2003).	Former Guilford Testing Coordinator fired	A district investigation that began in January 2003 revealed that the state end-of-course test questions were shared with students in some U.S. history and physical science classes before the tests were given. Three teachers and a central office testing coordinator resigned after district officials named them as participants.

This document is available on the Education Policy Studies Laboratory website at: **http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc**

Table 1, Continued

62. North Carolina	The Virginia-Pilot, Perry Parks (March 12, 1996).	Panel rules students' test scores are invalid	Writing test scores for 88 middle school students whose teacher gave them advance information about a statewide examination will be tossed out. Elizabeth City Middle School's overall average in state results will be hurt most by this invalidation of scores. According to the article, "The essay-style writing exam was supposed to be taken on Feb. 6 by every fourth-, sixth- and seventh-grader in North Carolina. But bad weather forced school closings and test postponements in several school districts, including Elizabeth City-Pasquotank. During the delay, information about the test question, or prompt, spread by word of mouth among some students and teachers who knew people in other districts. One seventh-grade teacher at the middle school, officials said, used that information to give students hints about the topic and format of the test on the day before Elizabeth City's exam date, Feb. 13. The question asked students to argue whether physical education should be a required class. This test, among others in the state, is used to empower local districts to hold teachers accountable for teaching students effectively." "State officials use the results of these tests to decide if teachers and principals are doing their jobs. That puts a lot of pressure on educators to come up with good numbers," officials said.
63. Arkansas	Arkansas Democrat-Gazette (April 10, 1990).	Teachers could be cheating: Change questions to keep integrity of test, analyst says	Arkansas's high-stakes standardized testing may be pressuring teachers to cheat. "We received a lot of letters from Arkansas teachers who didn't feel right about cheating" on state-mandated tests, Dr. John Jacob Cannell said in a telephone interview. Cannell said apparently some teachers are taking tests home and changing wrong answers to right ones and that because the tests have not been changed in several years, some teachers are teaching the tests to their students.
64. Louisiana	The Advocate (September 18, 1997).	Teacher admits fixing test answers	A Prescott Middle School teacher resigned after admitting to correcting answers on seventh-graders' standardized tests. Prescott Middle is one of 19 schools in the state where officials are calling for further analysis of multiple-choice LEAP tests because so many answers were erased and changed, according to the state Department of Education. According to Rebecca Christian, director of the state Education Department's bureau of pupil accountability, "The average number of erasures statewide is one per student. The average is then used to calculate a standard for flagging. The standard for seventh graders is eight erasures per student taking the test. Approximately 55,000 seventh-graders took the 78-question test."
65. Tampa, Florida	Tampa Tribune, Lynn Porter (February 21, 2003), p. 1.	Teacher may lose job over pretest	Reports on the behavior of one teacher as helping his emotionally handicapped students take a reading exam. The article alleges that he, "reviewed student answers, and if they were wrong, directed students to go back and try again."

This document is available on the Education Policy Studies Laboratory website at: **http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc**

Table 1, Continued

66. Maryland	Washington Post, Jay Mathews and Ylan Q. Mui (October 10, 2003), p. B01.	MD faults U.S. law for poor scores	Thirty Maryland elementary schools failed to meet NCLB requirements for adequate yearly progress because educators assisted thousands of children who have disabilities or limited English skills on a key standardized test. Article raises key concerns over how much to help/not help students with disabilities and students who are non-English speakers.
67. Florida	St. Petersburg Times, Robert King (May 9, 2002), p.1.	Test helper won't face any charges	Story about allegations against one teacher who acknowledged helping two students out on the statewide FCAT exam by pointing out problems they needed to rework. Prosecutors decided not to pursue criminal charges.
68. Oklahoma	Associated Press (April 10, 2002).	Teacher says she didn't break test security rules	A teacher had denied allegations that she helped her fifth-grade students cheat on state-mandated tests. "The state Education Department and the Oklahoma City School District accuse JoAnn Harding of allowing students to correct wrong answers. She's also accused of reading the math portion of the test, giving the test after the specified dates and not having an approved monitor for students taking parts of the test they missed."
69. Florida	Associated Press, The Bradenton (March 8, 2004).	Teacher accused of helping to cheat on FCAT	A Broward County teacher may be fired after 19 students alleged he gave them answers to last year's Florida Comprehensive Assessment Test and told them to keep it secret. The school superintendent is urging the local school board to fire the fifth-grade teacher for "misconduct, immorality, and moral turpitude." The teacher originally told officials that he "guided students to the right answer," and made it possible for students to "change answers." The teacher later said that students were lying about the allegation because they complained the teacher gave them too much homework. However, according to the story, one student claimed that the teacher had been doing this for five years and that "he had never gotten in trouble and nobody had ever told." This incident is the fourth FCAT cheating allegation in Broward County.
70.Ohio	U.S. News and World Report, C. Kleiner (June 12, 2000).	Test case: Now the principal's cheating	Reports on a story where one educator was accused of physically moving a student's pencil-holding hand to the correct answer on a multiple-choice question.

This document is available on the Education Policy Studies Laboratory website at: **http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc**

Table 1, Continued

71. Ohio and National perspective	The Times Union, Jay Mathews (June 4, 2000).	Cheating cases rise with test mandates	Fifth-grade teacher Barbara McCarroll was already puzzled and a little upset about her students' low test scores when her boss at Eastgate Elementary in Columbus, Ohio, approached her. How was it, the principal snapped, that the same children had done so much better on standardized exams the year before? After eight years of teaching, McCarroll knew it paid to be frank with children, so she put the question to them. She was not prepared for the answer: '"Well, Ms. McCarroll, that's because they gave us the answers and you didn't." Five months later, McCarroll sits at home, on disability leave since developing sleeping and eating problems. She says she was forced out of the school because she complained about coaching by teachers who administered last year's state proficiency test. She, her principal and her school join dozens of others across the country caught up in a rash of alleged cheating incidents seemingly brought on by a political movement to raise student achievement. At a time when superintendents are under pressure to increase test scores and hold principals and teachers accountable for student achievement, talk of cheating dominates the conversation in education circles. '"What we are seeing is what comes from the pressures of these high-stakes tests,"' said Vincent L. Ferrandino, executive director of the National Association of Elementary School Principals, who saw a similar scandal in Fairfield, Conn., while he was state education commissioner. '"There needs to be some discussion of the validity of this kind of assessment.'"
72. Florida	The Sun-Sentinel Jamie Malernee (Education writer), (March 17, 2004).	Parents, students protest Lauderdale teacher's removal over FCAT incident	The writer describes an alleged incident where one elementary school teacher was accused of helping students to cheat on the Florida statewide assessment, the FCAT. However, parents and students were rallying to the teacher's defense. "She is not the type of teacher who would help students cheat on the FCAT," according to more than 50 parents and students who marched outside the school on Tuesday, March 16. Parents and teachers were walking to support the teacher and to protest her removal from the "A-rated Ft. Lauderdale Montessori school." According to one 11-year-old, "She is the best teacher I've ever had." This student was in the class when the teacher was alleged to have told students they had wrong answers. According to this student, "Nothing happened. Somebody finished, she went through their test and she said, 'You have plenty of time, you should go through your work.' If what she did was wrong, every teacher should be out. They all say that." However, other students and the test proctor have alleged that this teacher gave students correct answers thereby forcing the director of assessment for the district to invalidate the scores of more than 20 fifth graders. The writer notes: "Whatever occurred, several parents at the protest said they don't think it merits pulling Bruening, who taught a mixed class of fourth and fifth graders, from the classroom while the investigation is going on." One parent said, "It's ridiculous, criminal, to take out a teacher who is such an asset to the school." One other retired teacher commented in the following way, "Teachers are guilty now until they are proven innocent. They don't make (the testing rules) clear. There's a lot of room for interpretation."

This document is available on the Education Policy Studies Laboratory website at: **http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc**

Table 1, Continued

73. California	Ventura County Star, Jean Cowden Moore (May 15, 2001).	2 schools face penalty in test-score blunder	Two Simi Valley high schools could have their test scores invalidated and find themselves knocked out of state rankings because they improperly prepared students for the Stanford 9, a statewide exam that most schools took this month. The schools, Simi Valley and Royal, also will likely not be eligible for any reward given to schools that improve from last year. That could be as much as $675,000 total. And freshmen, sophomores and juniors could lose their chance to earn $1,000 scholarships that will be awarded to students who score in the top 10 percent of their high school, or the top 5 percent across the state. Both schools gave students questions from past versions of the Stanford 9 to practice for this year's exam, said Kathryn Scroggin, assistant superintendent for the Simi Valley Unified School District. Some of those questions were repeated on this year's exam.
74. New Mexico	Albuquerque Journal, Andrea Schoellkopf (June 12, 2001).	Test breaches linked to teacher training	The state is investigating four school districts for allegedly breaching test security procedures at seven New Mexico Schools. These schools were accused of using test materials for test prep activities. "The districts are not following through with appropriate training," said Cheryl Brown-Kovacic, director of assessment and evaluation for the state Department of Education. "That was the big issue over and over again. They were using the test as prep material and saying they didn't realize what it was." The state announced Monday it had concluded its testing investigation into the four school districts Albuquerque, Gadsden, Gallup-McKinley and Santa Fe. The cases might be used to illustrate the problems that are coming about because of high-stakes testing, Brown-Kovacic said. The state had earlier blamed the problems on the new high-stakes approach taken to testing, where schools are placed on probation if they don't meet minimum requirements. Six of the seven schools that breeched security are rated probationary or just above because of performance on past tests.
75. Virginia	The Virginian-Pilot, Mathew Bowers (August 8, 2001).	VA Group reviews range of SOLS concerns	A legislative study group is thinking about a set of recommendations that would help to thwart cheating on the SOLS. Among the issues commission members raised are: Requiring that the same tests be given on the same days statewide, so students or teachers can't share the contents. Not repeating tests in the same year at the same school. Allowing teachers to see only the same day's test – and not those for the rest of the testing week. Recommendations are a result of several cheating allegations: Two other Chesapeake teachers resigned and a fifth was suspended with pay pending further action after other cheating allegations last spring. The incidents made up the bulk of 10 test security problems reported statewide out of some 2 million tests given. Test security is not a big problem, state officials told the commission. Two Oscar F. Smith High School history teachers face firing over allegations that they unfairly helped their students prepare for the state tests last spring. Apparently they reviewed material as the school division and state suggested, mirroring SOL formats and language. The state said, however, that they went too far, with their reviews markedly similar to the test given. Retesting was ordered.

This document is available on the Education Policy Studies Laboratory website at: **http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc**

Table 1, Continued

76. Louisiana	The Advocate, Charles Lussier (May 1, 2003).	Educators probe test allegations	A team of educators is under investigation for whether a teacher at Capitol High School breached test security by discussing the content of one section of a high-school exit exam administered in March 2003. The investigation began Tuesday during a routine review of the school's records, Baird said. The review was conducted by Capitol High's district assistance team. These are a group of educators not affiliated with the school who review its improvement efforts.
77. Buffalo, New York	Buffalo News, Darryl Campagna (December 17, 1999), p. 1A.	School board rehires teacher accused of cheating	A teacher resigned from teaching on the heels of allegations that he shared test questions with students before an exam.
78. Deltona, Florida	Associated Press (July 27, 2001).	Deltona teacher faces firing for helping students cheat on test	A teacher admitted helping high school students cheat on the graduation test. Allegedly, the teacher told a group of students, "Today, you all are going to pass high school," as she gave them the answers to the test.
79. Tampa, Florida	St. Petersburg Times, Editorial (March 11, 2000), p. 14A.	Raining dogs and FCATs	Two teachers were accused of giving students answers on the FCAT in April 2000. Editorial laments the pressures of standardized tests saying: "The FCAT is drowning schools, and for a simple reason. It is a political prescription for change, and it is being used in ways for which such tests were never intended."
80. Winston-Salem, North Carolina	Associated Press (June 8, 2000).	Teacher suspended after being accused of test tampering	A social studies teacher was suspended without pay after being accused of tampering with students' end-of-course tests. There were allegations that the answer sheets were tampered with.

This document is available on the Education Policy Studies Laboratory website at: **http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc**

Table 1, Continued

81. Memphis, Tennessee	The Commercial Appeal, Aimee Edmonson (Staff) (September 21, 2003).	Exams test educator integrity—emphasis on scores can lead to cheating, teacher survey finds	Report on many teachers who feel pressure to cheat to help their students. Teachers report that many of their colleagues do things to "help" their students on standardized tests. According to the article, cheating happens, although it is the exception rather than the rule, according to a survey conducted by The Commercial Appeal of 11,000 city and county public school teachers. Of 1,433 teachers and administrators who responded, almost nine percent said they'd witnessed impropriety on the part of teachers or principals during the Tennessee Comprehensive Assessment Program (TCAP) or Gateway exams. About 11 percent of city teachers said they had witnessed teachers or principals cheating; about three percent of Shelby County teachers said so. In 124 instances, teachers outlined a laundry list of rule-breaking: • Weak students herded to the school library to watch movies all week while others take the TCAP. • Teachers casually pointing to wrong answers as students work with their test booklets and answer sheets, admonishing them with such warnings as "you know better than that." • School counselors locking themselves behind closed doors after the test is done to erase "stray marks" on students' bubbled-in answer sheets. • Suspending a few borderline students for infractions just before test week so they can't bog down overall averages. • Some teachers even tattled on themselves. "There are never any monitors in the classroom during testing," said a city schools second-grade teacher. "Teachers can just close and lock the door and do what it takes to get those scores up!" Most respondents, however, said they had no direct knowledge of instances of teacher cheating and many insisted that it doesn't happen at their schools. "I work with professional educators who would never consider compromising themselves in such a manner," said one city high school teacher. Said a county middle school teacher: "Test materials are under tighter security than President Bush!" Pressure on teachers and principals to deliver good test scores has never been as intense.

This document is available on the Education Policy Studies Laboratory website at: **http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc**

Table 1, Continued

82. Texas	Shaila Dewan, Houston Press (February 25, 1999).	The fix is in: Are educators cheating on TAAS? Is anyone going to stop them?	Discusses how the results of the most recent TAAS raised suspicions about cheating. In one school, the results were almost "too good to be true: Every single fourth grader passed both the reading and math portions of the test." These results at Kashmere Gardens Elementary School raised speculation among the school's past and current teachers who suspected cheating was involved in the school's success that year. The article notes that teachers have incentives to cheat, "performing well has its rewards. Teachers receive candy, flowers – and pay bonuses. There are banquets and press conferences, banners and marching bands, most important in terms of public perception, school ratings: 'exemplary'…for schools that show muscle." Teachers who left the school remember being encouraged to "help students" do well on the test. And, analyses of erasures showed an unusually high number of erasures on student test booklets. For example, in 1997, one of Scott Elementary fifth-grade classes had a whopping eight erasures per child and "96 percent of them went from wrong to right – but it doesn't tell you how or why." According to one teacher, who left Kashmere, teachers were cheating as far back as 1994.
83. Las Vegas, Nevada	Las Vegas Sun Emily Richmond (January 8, 2004).	Teachers may have provided peek at exam	Two high school teachers are being investigated for allegedly distributing contents of statewide math proficiency exam to students before they took test. Teachers were accused of copying test questions by hand and then passing them out to students.

This document is available on the Education Policy Studies Laboratory website at: **http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc**

The first 22 articles in Table 1 document cheating by principals, counselors and others, sometimes alone, sometimes in collusion with teachers. The remainder of the table then presents dozens of stories about teachers' involvement in cheating. This large table unequivocally makes the case that something objectionable is happening in education. Stories arise from New York, California, Florida, Kansas, Utah, North Carolina, South Carolina, Pennsylvania, Massachusetts, Arizona and a dozen more states. Where high-stakes testing pressure is strong, and the press is active, we find many stories of cheating.

Instead of treating the incidence of the violations noted in Table 1 as a sign of an epidemic, we have simple condemnations of those who succumb to the moral pathogen, while the federal government and state educational agencies ignore the conditions under which the disease flourishes. The conditions fostering the corruption of teachers and administrators should receive more attention.

We focus in Table 1 on occurrences, but some work on rate or frequency of cheating has also been done. In spring 2002, in Chicago, economists Brian Jacob and Steven Levitt identified a set of classes whose scores on the Iowa Test of Basic Skills seemed unusually high and whose students showed a pattern of answers that seemed statistically unlikely to occur.[41] Those classes were retested. They showed a decline on the retest of a full year's grade-equivalent in scores. These researchers suggest that 3 to 5 percent of the teachers in Chicago were cheating.

In another study done by a team from Boston College, and more national in scope, we learn that violations of standardization procedures are common.[42] These violations, however, often seem to be different from what we simply call deliberate cheating.

This document is available on the Education Policy Studies Laboratory website at:
http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc

Nevertheless, the result is exactly the same. Violations of standardization procedures, like deliberate acts of cheating, yield scores of questionable validity.

The national survey by Pedulla et al.[43] revealed that about 10 percent of the teachers provided hints about the answers during the test administration; about 10 percent of the teachers pointed out mismarked items by students; about 15 percent of the teachers provided more time for the test than was allowed; and about 5 percent of the teachers engaged in instruction during the test. The survey also revealed, in what must be a rather large underestimate of the true state of affairs, that about 1.5 percent of the teachers actually admitted that they changed student's answers on the test. The immorality of teachers is often less self-serving than that of the Enron, Tyco, and J. P. Morgan executives, since many of these teachers have as their motive a desire for their students to do well on the tests. Thus some might be tempted to look the other way. But whether their behavior is judged to be either noble or ignoble, what is clear is that it decreases the validity of the tests and adds uncertainty to the decisions that are made on the basis of test scores.

Student Cheating and the Inevitability of Cheating When the Stakes are High

A recent movie titled The Perfect Score, shows how important college entrance tests can be in the lives of some students. In the movie, the students set out to steal the test. They justify their theft because they know that the test should not define who they are or what they will be, and they are sure that the test cannot assess what they actually know. They resent the importance of the test in their lives and many in the audience

This document is available on the Education Policy Studies Laboratory website at:
http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc

become sympathetic to their argument. But these fictional students have their real life counterparts who are described in Table 2, where we have collected articles that document student cheating and instances of students speaking out against the test.

Although it is likely that a student cheated on the day that the first test in the history of schooling was announced, it was not likely that such widespread cynicism about the test was present on that day. The articles in this table reveal that huge percentages of school children have cheated on tests (see Article 2). Perhaps more surprising is that the rate for cheating by students in religious schools may even exceed that of the students in public schools (see Article 3)!

This document is available on the Education Policy Studies Laboratory website at:
http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc

Table 2: Student Cheating and the Inevitability of Cheating When the Stakes are High

Location of Story	Source	Headline	Story
1. Utah	Salt Lake Tribune, Katherine Kapos (February 8, 2000).	State tries to head off cheating: More tests mean more pressure on schools, more temptation to cheat	In recognition of the growing pressure on students and teachers to do well on standardized tests, the state has proposed a set of procedures for trying to thwart cheating. According to the article, legislators recognize that the cheating is more likely when stakes attached to test scores are high. Under the proposed policy, it would be up to each district to secure tests in an area where only authorized personnel have access, and to impose security measures before and after testing periods. The State Office of Education would perform periodic audits to ensure that district procedures guarantee that tests are secure and properly accounted for. School district employees and personnel would be forbidden from copying or reproducing the tests. The policy also outlines when tests will be given and how the results will be collected by the state for public reporting.
2. New York	New York Post, Carl Campanile & Jennifer Fermino (May 12, 2004).	Kids cheating their way through school	A Gallup poll revealed that nearly half of the nation's teenagers admit that they cheat on their school exam. Based on interviews with 2,000 students (run on the internet), the survey found that 46 percent of students said they cheated at least once on a test or exam. According to the article, "Some city high school kids interviewed yesterday scoffed at the poll results, insisting that cheating is even more pervasive in Big Apple schools. 'Only 46 percent? Everybody cheats,' boasted Peter, a 16-year-old junior at the La Guardia HS for Performing Arts on the Upper West Side. Rachel, a ninth grader at La Guardia, said she's been cheating since middle school. She claimed she's even stolen copies of tests before they were administered. She said she's looked over her smarter classmates' shoulders as well. Frank Newport, editor-in-chief of the Gallup Poll, noted that 'the federal No Child Left Behind Law requires school districts to test students in grades three to 12. New York has among the most stringent graduation requirements in the nation, and kids fearing failure may cheat,' Newport said."
3. New Orleans, Louisiana and National perspective	Times -Picayune, Adelle Banks (October 26, 2002), p. 13.	Student cheating appears on rise: Religious school kids lead, survey shows	A report by the Josephson Institute for Ethics found that 78 percent of students attending private religious schools said they had cheated at lest once on exams in the past year, compared with 72 percent of students at other schools.
4. Oakland, California	Op-Ed.	Cheating is now disturbing norm.	Editorial lamenting the fact that student cheating has become rampant.

This document is available on the Education Policy Studies Laboratory website at: http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc

Table 2, continued

5. Georgia	The Augusta Chronicle, Kamille Bostick (Staff Writer) (December 16, 2003), p. D3.	Cheat tricks/Failing grades: Punishments are in store for students who get caught	Story uncovers various ways students cheat on quizzes, exams or vocabulary tests. Marissa Harding, 18, a senior at Cross Creek High School, has seen students take cheat sheets into vocabulary quizzes, sometimes even putting the answers on the back of chairs they sit behind. "They spend a lot of time thinking of ways to cheat rather than trying to study," Marissa said. "I don't know if they don't want to learn it or like to see if they could get away with it." Many students cheat because of the increased pressures to do well. One senior said, "With so much emphasis on grades and so much riding on academic excellence, students resort to cheating. … There's so much pressure on students today – especially upperclassmen – they start to think: 'I want to pass this class; I want to get into Georgia; I need a good grade,'" she said. "The pressure is overwhelming sometimes."
6. California	The Mercury News, Kamika Dunlap (Staff Writer), (January 30, 2004).	200 parents pack meeting on cheating at Saratoga High School	A meeting of educators at Saratoga High School with over 200 parents discussed concerns about increased student cheating. According to the principal, "I think there is a great pressure on our kids, and we want them to be successful and have material rewards," Skelly said. "But, I think what we really want for them is to have strength and character to handle whatever comes their way." A parent was also concerned, "We have raised the bar up, and kids are very nervous and anxious," said Steve Tilles, father of a Saratoga senior. "I think competition for grades is extreme for all students and that grade inflation has raised anxiety. "The meeting was held to address recent news that a small group of students at the school allegedly used tiny computer devices to find teacher passwords and steal English department tests and answers. Expulsion hearings for two of the students involved begin today at the Los Gatos-Saratoga high school district. Three other students are being investigated, one for allegedly breaking into a school computer and trying to change a grade, and two for allegedly stealing a printed test and saved electronic copies.
7. Columbus, Ohio	Columbus Dispatch, Ruth E. Sternberg (October 27, 2003), p. 1B.	Schools tackle test security	Ohio teachers are prepped to look for student cheating. This is on the heels of suspicions of student cheating from the spring 2003 semester, when 90 students' math scores suddenly shot up. Officials suspected "someone had sold them a test, and that they had sneaked in their answers."
8. Washington, DC	Washington Post, Brigid Schulte (November 26, 2002), p. B03.	SAT firm probes cheating at Landon	ETS officials investigated allegations that eight students in a Bethesda high school admitted to copying and sharing answers to boost their SAT scores.

This document is available on the Education Policy Studies Laboratory website at: **http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc**

Table 2, continued

9. Washington, DC	Washington Post, Liz Seymour & William Branigin (June 2, 2001), p. B03.	Theft forces school to give state test again	Tenth graders from a Fairfax County high school were accused of stealing a copy of the chemistry test two days before they were to take it, forcing upwards of 250 students to retake the test.
10. Texas	The Houston Chronicle, Melanie Markley (December 18, 2002), p. 40.	Districts warned of TAKS cheating: Possible cases found in early test	In Texas schools, 46 students were suspected of cheating on a writing test that was given to prepare students for the new state exam. Testing scorers contacted 18 school districts and two charter schools that some of the writing samples showed evidence of plagiarism.
11. Milwaukee, Wisconsin	Milwaukee Journal Sentinel, Nadia Lerner (April 27, 2003), p. 1L.	Cheating on their minds	Article on why more students are engaging in cheating.
12. Spokane, Washington	Spokesman Review, Staff (March 31, 2003), p. B1.	Cheating ways: Many students are willing to sacrifice integrity for good grades	Story talking about the prevalence of cheating in Washington schools. Story quotes several students and their perspectives on cheating. One student said that "everybody cheats." One student said that she cheated "because I had the chance, basically."
13. Boston, Massachusetts	L. Seymour; W. Branigin	Theft forces school to give state test again	In Boston, three 10th graders were accused of stealing a copy of the standardized chemistry test two days before they were to take it, causing more than 250 students to retake it.
14. Westport, Connecticut	New York Times, Jane Gross (November 26, 2003).	Exposing the cheat sheet, with the students' aid	Story highlighting the ways students cheat in response to the unrelenting pressure to succeed. According to the article, one student had cheated by printing the answers on the back of a water bottle label. According to the story that reports on a discussion one principal had with his students, "the students used the word 'contagious'… if they knew the kid next to them was doing it, and winding up with a higher grade-point average, it was difficult not to participate. If you look at it a certain way, it's a reasonable response to a set of unreasonable expectations. But, the students told me they'd had enough of it. They want it to stop. They need adults to take it seriously."

This document is available on the Education Policy Studies Laboratory website at: **http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc**

Table 2, continued

15. Nevada	Las Vegas Sun, Associated Press, (January 18, 2004).	Student testing irregularities increase in Nevada	At Spanish Springs High School in Sparks, Nevada, a student wrote answers to the math portion of the test on a separate sheet of paper, then provided them to another student during a prearranged restroom visit. Both students were suspended and their tests were invalidated.
16. National perspective	Cox News Service, Rodney Ho (February 1, 2004).	SAT perfectionism can tempt students to cheat their way to 1600	Talks about, "The Perfect Score," a recent movie about students who conspire to steal a copy of the SAT test from ETS in Princeton. Raises questions about real instances of student cheating on the exam. ETS acknowledges that about 300 tests a year are thrown out for cheating. The article points out several instances of student cheating. • Thanks to a lax proctor, eight boys in a Bethesda, Maryland prep school admitted to copying and sharing SAT answers in 2002 • A dozen students used cell phones to exchange answers via text messaging and downloaded information from the Internet in Rockville, Maryland, last year. • One high school junior in GA said she knew someone who hired somebody else to take the test for them three years ago and got away with it. West coast students got answers from east coast students, thanks to a three hour time difference. Others have used different colored M&M's to signal answers – and many go back to earlier sections during breaks (a violation of SAT rules).
17. Ohio	Columbus Dispatch, Editorial & Comment (November 8, 2003), p. 10A.	Overdue prevention: Columbus officials should crack down hard on students who try to cheat on tests	The Columbus Board of Education should approve a new, detailed test-security policy that clarifies what teachers should look for and spells out suspicious signs that could signal cheating, such as excessive eraser marks on answer sheets. According to the article, a national survey last year found that nearly three-quarters of 12,000 high school students had cheated on an exam at least once in the past year.
18. Cleveland, Ohio	The Plain Dealer, Opinion (July 14, 2003), p. B6.	Test cheaters shouldn't prosper	Allegations that 100 students cheated on the state proficiency test went nowhere after proof and confessions were hard to come by. Only one student was seriously investigated, but no students were punished. Students interviewed told officials about rumors that a tutor at a city recreation center was selling test answers. Teachers said that large test score increases on the ninth grade math proficiency test made them suspicious. Many of the students with such surprising results took the test on a make-up day. The article claims that the state's unwillingness to punish students only further tempts students to cheat. This outcome sends the wrong message to struggling students. Can't pass the test on your own? Get the answers from someone else, make sure you both keep your mouths shut and waltz right on toward graduation.

This document is available on the Education Policy Studies Laboratory website at: **http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc**

Table 2, continued

19. New Mexico	Online: KOBTV.com (February 7, 2004).	APS teachers ask district to restrict student cell phones	There is a national trend of students using cell phones as a means of cheating. In New Mexico, teachers are responding to this by asking the district to restrict students' use of cell phones. Some students there had been caught "using text-messaging features to cheat on tests." High school student Desean Smith says text messaging "gives you a way to talk to somebody that's not there without teachers knowing." Amanda Black, another high school student, says secret text messaging is quite easy. "After you do it for a while, you get the hang of it to where you don't even have to look at your phone. You can do it and still take notes and listen to the teacher," she said.
20. California	Knight Ridder Newspapers, Suzanne Pardington (February 25, 2004).	Students cheating with cameras, text-messaging on cell phones	Story about the growing concern over the use of technology to help students cheat. Specifically, with the technology of cell phones and text messaging, more teachers are asked to be more aware of their use during testing situations.
21. Lake Tahoe, California	The Daily Tribune, William Ferchland (April 30, 2004).	Copy Machines	Story about the various ways students are using technology to cheat. "Nowadays students are able to manipulate technology and traditional methods which have forced teachers to battle back with the Internet and philosophy. … Some say the pressure to succeed has increased with colleges accepting fewer candidates, and tests increasing in importance. Others believe the availability of teachers, make-up opportunities and summer school has relieved the need to cheat." In South Tahoe High School, newly formed discipline committee is considering revamping its cheating policy.
22. Florida	Sun-Sentinel, Jamie Malernee (June 8, 2004).	100 Broward students under investigation for possible FCAT cheating	More than 100 students' test scores were identified as having unusual patterns by a computer program the state uses to search for signs of cheating. In Florida, students have to pass the FCAT to receive a diploma, thus this investigation is holding up at least 100 students' diplomas. Students' scores rose so much that they were flagged as suspicious and their scores were held back for further review into allegations that they had cheated. But, teachers and principals have come forward insisting that many of these students had not cheated and that they shouldn't be punished for making dramatic improvements. Yet Cornelia Orr, head of assessment for the Florida Department of Education, maintains that the main reason scores are reviewed is a hard one to refute: students having the same answers on their score keys. Still, many students feel they are being unfairly targeted and punished. One student, Alvaro Plazas was one of those suspected for cheating. And at least one teacher vouched for him: "I would stake my teaching certificate on him not cheating," said Infinger, who wrote to Tallahassee on behalf of Plazas. "This is my 15th year teaching and he would have to be in my Top 10 students in terms of being responsible, polite and hardworking."

This document is available on the Education Policy Studies Laboratory website at: **http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc**

Table 2, continued

23. Florida	Miami Herald, Steve Harrison (February 24, 2004).	State to check closely for FCAT cheating	After a Herald analysis revealed the possibility that dozens of schools statewide had posted FCAT results that didn't seem to make sense, some test sheets will be examined by hand to look for erased marks and other marks that might signal cheating. According to the report, Broward and Miami-Dade had a larger number of schools posting statistically unlikely results on the Florida Comprehensive Assessment Test than other school districts. Some had already been accused or investigated for cheating, but many had not. Although resources typically are thin for this type of probe, one spokesperson said that more will be done to investigate cheating allegations and concerns.
24. Louisiana	Times-Picayune, Chris Gray (June 1, 2001).	21 face cheating charge at Landry: Written responses on LEAP identical	At least 21 students at Landry High School have been accused of cheating on this year's LEAP and must attend summer school and pass a retest to be promoted. "Their scores were voided due to plagiarism," state testing director Scott Norton said. In total, the scores from 21 math tests and 14 science tests at the school were thrown out because students gave identical responses on the written portions of the exams, he said. This isn't the first time state educators have accused students of cheating on the Louisiana Educational Assessment Program test; last year about 100 of the exams statewide were discarded because they seemed suspicious, Norton said. But students who have cheated usually are caught in groups of two or three; not 21. Seven other schools in New Orleans also have had scores zeroed-out because of plagiarism, Norton said, including 11 fourth-grade science tests at Robert R. Morton Elementary School and eight eighth-grade social studies tests at James M. Singleton Charter Middle School.
25. Louisiana	Times-Picayune, Chris Gray (June 30, 2001).	Students accused of exam cheating: Test official -- Teacher may be involved	As many as 37 sophomores at Joseph S. Clark Senior High School are accused of cheating on this year's Graduate Exit Exam – probably with the aid of a teacher. The students are accused of having strikingly similar answers on the written portion of the exam. When more than 10 students in one classroom have identical responses to the questions, the state Department of Education alerts the district that a teacher may have been involved in the cheating, state testing director Scott Norton said. "It's a flag that it might be more systematic," he said. In a memo to the Orleans Parish School Board in early June, the state listed 29 students who had similar responses on the math portion of the GEE, which students must pass to get a diploma. Three other students are accused of collaborating on the English test and five more posted suspicious answers on both sections of the exam. A similar situation occurred at John McDonogh High School, where nine tenth-graders posted similar answers on their English tests and four possibly plagiarized their math tests. Incidents involving smaller groups of tenth graders also took place at Abramson, Landry, Carver and Lawless high schools.

This document is available on the Education Policy Studies Laboratory website at: **http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc**

Table 2, continued

26. Houston, Texas	The Houston Chronicle Sherry Thomas (September 26, 2002), p. 1.	Who's cheating? Academic dishonesty gets educators' attention	A Houston school district pairs up with an Internet company to be able to analyze papers for plagiarism.

This document is available on the Education Policy Studies Laboratory website at: **http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc**

This table also highlights two trends that seem noteworthy. First is the increasing use of technology in cheating. An older teaching and administrative work force is not always aware of the latest in sophisticated messaging capability with which a younger generation is familiar. Some ingenious forms of cheating are described in Articles 19 - 21. The problem is worldwide.[44] Second, and by far the most serious issue that emerges from studying this table is that we have designed a system that expects our children to cheat. As seen in Article 1 and others throughout this table, there appears to be the assumption that cheating is inevitable; an ordinary activity under the conditions of schooling that our nation has designed. In response to the incidents of cheating, of course, schools and state departments of education take steps to impede their occurrence. It is not clear how much time and money is being invested in policies and technology to detect and thwart cheating, but it is surely substantial. Yet, reactive approaches to controlling behavior seldom work.[45] As building more prisons does nothing to change criminal behavior, it is likely that better detection and prosecution for cheating will not deter it. Instead, we need to design schooling practices that help us raise children less tempted to transgress.

Reactive approaches to controlling behavior seldom work. As building more prisons does nothing to change criminal behavior, it is likely that better detection and prosecution for cheating will not deter it.

Campbell's law can help us to predict when corruption of individuals is likely to occur. That analysis should be the basis for a national dialog about the design of school practices that promote moral development and helps to curb school practices that compromise those who attend and work in our schools. Through dialog we can ponder

This document is available on the Education Policy Studies Laboratory website at:
http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc

one of Einstein's remarks, tailored perfectly to our high-stakes testing era: "Not everything that counts can be counted, and not everything that can be counted counts."

Excluding Students from the Test

If students do not finish high school their chances for an economically productive life are seriously compromised. Chicago, one of many urban areas with high dropout rates, reports that 22 percent of all residents between 16 and 22 are out of work and out of school. That constitutes 100,000 people in Chicago. We also have 5.5 million out of school and out of work young people in the U.S.[46] Such youth were once metaphorically called "social dynamite," energetic, unstable, angry, and ready to explode. So it is always in the best interests of society to keep students in school and to help them to get degrees. High-stakes testing, however, has exacerbated the traditional problem of keeping disaffected youth of all ability levels in school.

Abandonment of our youths, many of whom can be helped to finish school, has reached epidemic proportions as a result of high-stakes testing. It is widely known that if the disaffected or weakest students can be pushed out of schools or allowed to drop out, then the test scores at the school or districts that lose these students will go up. As stories in Table 3 illustrate, many students are being left behind simply because they are the lower scoring group, a group that includes predominantly poor and minority youth, youth with special needs, and those for whom English is a second language.

While the departure of "weaker" students raises test scores, their rates of high-school completion are also monitored and become another indicator of school success (or failure) for a school or a district. So Campbell's law predicts that dropout rates and high

This document is available on the Education Policy Studies Laboratory website at:
http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc

school completion rates will be corrupted, too. And they are! School districts have simply lied about their dropouts and their completion rates or so fudged the data that it is uninterpretable. In this way three indictors are corrupted: drop out rates, high school completion rates, and achievement test scores. The difficulties of interpreting dropout rates is discussed in the next section, here we look primarily at the tendencies of schools to "push" students out and to stand by while their lower achievers give up and drop out.

This document is available on the Education Policy Studies Laboratory website at:
http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc

Table 3: Excluding Students From the Test

Location of Story	Source	Headline	Story
1. Tampa, Florida	Tampa Tribune, Ellen Gedalius (August 14, 2003), p. 1.	School allows student back after months of confusion	A school sends a letter home to a student encouraging him to dropout of school and informing him that "the traditional school setting is not suited to all students." The student had a low GPA and failed portions of the FCAT. But, even though he is eligible to take more courses to bring up his GPA and can retake the FCAT several times, the school tried to "push him out."
2. New York	New York Times, Tamar Lewin (September 15, 2003).	City to track why students leave school	New York City school officials are accused of pushing struggling students out of the school system. There are allegations that thousands of students, some as young as 16, have been pushed out of high school against their will. Chancellor Joel Klein acknowledged that there was a widespread problem that was a 'tragedy' for many students.
3. San Francisco, California	San Francisco Chronicle, Nanette Asimov (March 3, 2003), p. A1.	Disabled students call test unfair	Many disabled students are feeling the pressure of California's exit exam. The overall passing rate for the Class of 2004 was 48 percent whereas the passing rate of disabled students is 13 percent. Disabled students feel "pushed out" because they work hard, but cannot overcome the exit exam barrier to receive a diploma.
4. Connecticut	Hartford Courant, Robert A. Frahm (October 19, 2003), p. A1.	Big test, no hope: No Child Left Behind act offers no breaks for language barrier	Jose Torres was forced to take the statewide assessment even though he doesn't speak English. Principal was outraged at the insensitivity of NCLB requirements that mandate special education and non-English speaking students undergo the rigors of an exam they are destined to fail.
5. San Bernardino, California	San Bernardino Sun, Matt Bender, Selicia Kennedy-Ross and Tiffany Maleshefski (January 29, 2004).	Area schools stumble: Most fail to meet objectives put forth by No Child Left Behind	One district assistant superintendent worries about the effects of provisions in NCLB that would unfairly punish LEP students. The story notes that "Among his biggest concerns: Testing limited-English students when they've only been in the United States for a year or less, and expecting some special-education students to perform at grade level when their IQs suggest they can't." Students are also stressed and they offered mixed reactions to the focus on accountability. "A lot of kids don't like (standardized tests) because they have to take a test to graduate and they're like, 'But I'm taking a class',"said Erin Vetere, a senior at Pacific High School in San Bernardino. Other students said they felt they were being asked to hit too many targets every year. "It's just a lot of pressure," said Matthew Magdonadlo, a senior at San Andreas High School in San Bernardino.

This document is available on the Education Policy Studies Laboratory website at: http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc

Table 3, Continued

6. Virginia	The Virginian-Pilot, Nancy Young (November 10, 1999).	Some schools exclude students from SOLS: But efforts to skew scores are "rare" testing official says	In several local schools, the percentage of students who took SOL tests at the elementary and middle school levels was about 10 to 25 percentage points below the state average, according to figures from the state Department of Education. At the elementary school level, most of the students excluded are those with learning disabilities. At the middle school level, absenteeism comes into play. Both sets of students are more likely to bring down test scores. As more and more states turn to high-stakes testing, such as the SOLs, some fear that pressure will mount on schools to begin excluding students who may perform poorly. "It really depends on the teachers, especially in special education," said Alexander Saulsberry, principal at Newtown Road Elementary School in Virginia Beach. Students "might be age-appropriate socially for fifth grade, but they're functioning at a third-grade level. To have students where emotionally they're going to be zapped – why do that?" Saulsberry's school excluded more fifth graders than any other South Hampton Roads school on three of four SOL tests. Seventy percent of his fifth-graders were tested in history or science last year, down from more than 90 percent the year before. The school also posted impressive gains in test scores – both at the fifth-grade level where far fewer students were tested than in 1998 and at the third-grade level, which had comparable percentages tested in 1998.
7. Denver, Colorado	Rocky Mountain News, Kim Franke-Folstad, (Staff Writer) (October 12, 1997), p. 6A.	Standardized tests don't tell a school's whole story	Two days before the big test, the writer's son came home with a note from his former special education teacher (the son was currently in a regular education classroom) that said she didn't think Ben should participate in the statewide assessment of basic skill. The note stated that Ben wasn't doing as well as his classmates and it might hurt his 'self-esteem.' The article's author notes that the pressure on schools to produce student achievement improvement is forcing schools to keep students who might lower the school's results from taking the test.
8. Southern Pines, North Carolina	Associated Press (October 8, 2003).	Teachers say they were told to keep bad students from taking SAT test	Two English teachers at a North Carolina high school said they were told to dissuade low-performing students from taking the SAT. Three former students accused the former principal of discouraging them from taking the college admission exam.

This document is available on the Education Policy Studies Laboratory website at: **http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc**

Table 3, Continued

9. San Antonio, Texas	San Antonio Express News, Lucy Hood (August 17, 2003), p. 1A.	Lessons of change: The newest immigrants challenge the U.S. education system	Story about the plight of LEP students in Texas (and across the nation in general). The article talks about the law forcing LEP students to take the English version of the TAKS and the burdens that places on both the students and their teachers. "There are no figures that show the failure rate of immigrant students, but the scores of those classified as limited English proficient are far below the state average at the high school level." The article goes on to say that in Texas "the LEP 11th-grade passing rate, for example, was 15 percent in the spring of this year, compared with 49 percent for all students. For Anglos, it was 59 percent; for Hispanics, 38 percent; and for African Americans, 33 percent." "The testing stuff is really deadly," said Roger Rice, co-director of Multicultural Education, Training, and Advocacy, which frequently provides legal representation for immigrant students. "I see the testing push as something that is increasing dropouts, putting pressure on schools to push kids out, and making it more difficult for kids to get the kind of broad and solid education they would otherwise be looking for."
10. New York	New York Times, Tamar Lewin (September 15, 2003).	Education: The pushouts	In New York City, there is reason to believe that students who may bring down school-wide achievement averages are being "pushed out" of school before they reach their senior year. Few students are told they have a right to stay in school until they are 21. Instead, they are encouraged to enroll in an equivalency program.
11.Connecticut	Associated Press (April 1, 1999).	State finds East Hartford illegally exempted students from taking mastery tests	An investigation by the State Department of Education has found that East Hartford illegally exempted some low-achieving students from taking the Connecticut Mastery Test. It was found that during at least part of one year during the 1990s a policy was in effect in which any potential special education student was exempted from taking the test – a violation of state law.
12. Massachusetts	Masslive.com, Azell Murphy Cavaan (September 8, 2003).	MCAS refugees flock to GED	Increasing numbers of students from Springfield, Massachusetts are giving up earlier and opting for the GED. According to one student, "MCAS is the No. 1 reason … some students drop out of high school and get a GED just so they don't have to deal with the rigors of that test." And the rates are growing disproportionately among ethnic minority students – 50 percent of Latino students and 33 percent of African American students quit before their senior year.
13. Raleigh, North Carolina	Associated Press (July 13, 2000).	Critics say ABCs of accountability lead to increased suspensions	Figures from some school systems showed a spike in student suspensions during the last month of the school year, when teachers gave end-of-course and end-of-grade tests. As one educator was quoted, "What I hear is that they believe a lot of kids who aren't making it academically are becoming more vulnerable to becoming suspended and expelled so they don't weigh down the test scores."

This document is available on the Education Policy Studies Laboratory website at: **http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc**

Table 3, Continued

14. Massachusetts	The Boston Herald, Kevin Rothstein, (February 26, 2004).	Minority dropouts sky-high in Mass.	According to a Harvard report released on February 25, 2004, Hispanics in Massachusetts drop out of high school at the second highest rate in the country, and only half of the state's black students graduate. According to the article, "The study estimated that only 36.1 percent of Bay State Hispanics graduate, fairing only better than New York and falling far below the national rate of 53.2 percent. Researchers excluded 17 states from their Hispanic rankings because of not enough data. Black students in Massachusetts graduated at the 13th lowest rate in the country, excluding 11 states where the rate could not be calculated. Ranking the worst was New York, where only 35 percent of black students graduated." The article goes on to argue, "Some blame high-stakes tests for helping push students out of school. Pressure on schools to produce good standardized test results has led to documented cases in some states of students being pushed out of school, critics say."
15. Alaska	Anchorage Daily News, Katie Pesznecker (May 10, 2004).	Minorities struggle more with exit exam	According to the latest statewide test results, minority students in Alaska struggle with the state's exit exam more than white students, and a proportionately larger number of them will be denied diplomas as a result. In Alaska, the achievement gap has been evident for years, but this was the first year that 12th graders had to pass the graduation exam to get a diploma. Since 1998, and the exam's inception, minority students, particularly Native American students, have had a harder time passing the exam than their white counterparts. A similar trend exists for children from low-income families and students with limited English skills or learning disabilities. Sarah Scanlan, director of education with the nonprofit group First Alaskans Institute, said some students, faced with passing the test on a fifth and final attempt before graduation, are simply giving up and dropping out instead of being humiliated by another failure. "Who would have thought that we would have been causing more kids to drop out?" Scanlan said. "And when that's staring us in the face, the bureaucrats ought to be saying: 'Oops, back to the drawing board; let's see what we have to do to fix this unintended consequence.'"
16. Carthage, North Carolina	Associated Press (October 16, 2003).	Moore school board backs superintendent despite SAT bonuses	A superintendent was allowed to keep his job after an investigation that found that for two years, administrators at North Moore High School sought to improve SAT scores by shaping which students took the college-entrance exam. About $2,206 in money under the principal's discretion from drink machines and the student activity fund were used to reward selected students. Students with one of the school's best three scores earned up to $150 and the money was also used to pay the test-taking fees of some students. These apparently were students who might do well. The schools' own funds expended in this way were later reimbursed by the state from an account that was supposed to help lower achieving students.

This document is available on the Education Policy Studies Laboratory website at: **http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc**

Table 3, Continued

17.Florida	St. Petersburg Times, Rebecca Catalanello (Times Staff Writer), (May 9, 2004).	Experts: Retaining disabled students can breed failure	According to the article, third-graders with disabilities failed the state's standardized reading test at twice the rate of their non disabled peers – which means that the type of student most likely to be held back from fourth grade is a child "whose struggle to learn is caught up in a battle against physical impairment or learning disability." Preliminary figures from this year's assessment suggest that for the second year in a row, "thousands of Florida's third-graders are facing the fact that a low score on the reading test could prevent their promotion to the fourth grade. For 8,300 of them, this is the second year they have flunked the test. More than half of those students have a disability." Susan Rine, administrative assistant in Pasco Elementary Schools, "worries that holding exceptional students back a grade or two is more likely to frustrate them, thwart learning and ultimately foster poor citizenship. 'I just don't know how many times you can hit somebody over the head and say, You're not good at this, you're not good at this.'"
18. Washington	Tri City Herald, Brent Champaco (July 16, 2004).	Sunnyside petitions for WASL substitute	According to some, many students would rather quit school than try to pass the Washington Assessment of Student Learning (WASL). The state-mandated test, according to the article, drains confidence from kids who are struggling academically, especially Latinos – according to one Sunnyside School Board Member. And, officials predict that test anxiety will only worsen by 2006 when the WASL becomes a high school graduation requirement. It's the reason the Sunnyside School Board gave its OK in late June for Garza to craft a legislative proposal that calls for replacing the high-stakes WASL with another, less-stressful test. Frustration with the WASL prompted the grass-roots uprising in Sunnyside, where more than 80 percent of its 5,600 students are Latino. At Sunnyside High School, about 18 percent of students met WASL standards in math and 25 percent met standards in writing last year. A group called Nuestra Casa, composed of local Latino parents concerned with the WASL and other educational issues, pleaded with the school board to address the Legislature, officials said. Educational barriers like language and poverty make passing the comprehensive test impossible for many Latinos and non-English speakers, according to the "Latino/a Educational Achievement Project," an Issaquah-based advocacy group for student achievement.
19. Seattle, Washington	The Seattle Times, Jolayne Houtz and Linda Shaw (December 7, 2003), p. A1.	1 in 4 high-school students here fails to graduate on time	A new analysis shows that the percentage of students who make it through area high schools in Seattle, Washington, on time is significantly lower than had been thought, and some worry it might be getting worse. More than 300 freshmen took their seats in one area high school in the fall of 1998 and only 145 seniors left with diplomas. What happened to the other half of this class is unknown. Some moved or earned a GED instead. But, according to the article, it is a safe bet that a great many of them dropped out. Even more disturbing is the dramatic disparity in graduation rates among youth of various ethnic backgrounds. The on-time graduation rate for white students is 70 percent, 77 percent for Asian, fewer than 50 percent for Latino and 53 percent for African American students.

This document is available on the Education Policy Studies Laboratory website at: **http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc**

Table 3, Continued

20. Boston	Boston Globe Michele Kurtz (August 14, 2003), p. B7.	Tests seen increasing dropouts	Concern that the Massachusetts statewide achievement test (MCAS) is leading to higher dropout rates.
21. New York	New York Times, Diana Jean Schemo (August 30, 2004).	School achievement reports often exclude the disabled	The trend of excluding and avoiding accountability for the disabled population is alarming activists who see it as "an erosion of the education act's disclosure requirements. In them, parents and advocates say they saw a crucial level for helping their children meet higher academic standards, and a way of finding out which schools were meeting the challenge."
22. Massachusetts	The Daily Free Press, Steve Reilly (April 14, 2004).	MCAS may affect state dropout rates	Story reporting on a research study that found a relationship between the institution of high-stakes exit exams and student dropout rates. The report found that "9,289 ninth through 12th graders dropped out in 2002-2003, an increase to 3.3 percent. Seniors dropped out the most and freshmen dropped out the least." In the study of dropout rates across ten states, nine of the ten states with the highest dropout rates in 1986 had some form of exit or minimum competency exam. Thus, the report concludes the highest drop out rates are associated with higher stakes and less flexible standards than the states with the lowest dropout rates.
23. National perspective	Orfield, G., Losen, D., & Wald, J. (2004). Losing our future: How minority youth are being left behind by the graduation rate crisis. Cambridge, MA: The Civil Rights Project at Harvard University.	Research Report	Study investigating how minority youth are being left behind by the graduation rate crisis. As part of this report, the authors compiled a list of stories from students describing the ways in which they were "pushed" out of the test or out of school.
		Story 1: Alabama	In March 2000, 16-year-old Renae (pseudonym) enrolled in the Birmingham city schools adult education program. She pulled out a folded piece of paper to give to the instructor that said, "Withdrawal. Reason: Lack of Interest." Renae's mother later told the instructor that she had tried to get her daughter back in school but was informed that her daughter's standardized test scores were low and that she probably wouldn't graduate. Despite her mother's insistence that she wanted her daughter to remain in school, school officials refused to readmit her.

This document is available on the Education Policy Studies Laboratory website at: **http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc**

Table 3, Continued

		Story 2: Alabama	Of all the pushed out students who attended an adult education program, none had voluntarily withdrawn from their secondary school. In fact, some had actually returned to the school with their parents or guardians and asked to be readmitted, but their requests were denied. Parents had not been included in the withdrawal meetings, and some parents did not even know their children had been withdrawn. (p. 25).
		Story 3: Alabama	Birmingham city schools Board representatives have now admitted that 522 students were administratively (involuntarily) withdrawn in the Spring of 2000. To this day, students continue to be "withdrawn" from school for lack of interest, academic failure, and poor attendance, and the Birmingham schools continue to be under enormous pressure to raise standardized test scores. (p. 26).
		Story 4: Illinois	A number of recent Chicago "dropouts" report that they were pushed out of public high schools by officials who told them that their truancy or bad grades showed that they did not want to be there. Illinois law allows school officials to "dis-enroll" 16-year-olds who can't be expected to graduate by their 21st birthday. Seventeen-year-old Jennifer (pseudonym) said she started missing school regularly in her junior year because she was having problems with chemistry. "I felt stupid and I couldn't get the help I needed. Then one day I went to first period class and they told me I wasn't on the roster anymore. I was shocked." Another "dropout" John, also 17, said he started missing high school because of family problems, including having to stay home with an ill sibling. "I came back and they told me I wasn't on the roster anymore." (p. 38).
		Story 5: Florida	In July of this year, 17-year-old Danny (pseudonym), an African-American male between his junior and senior years at a high school in Tampa, Florida, received a letter saying that he could not come back to school in September. The reason given was that he had not passed the FCAT. Danny had, along with the entire junior class, taken the mandatory test in the spring. He had been unaware, until receipt of this letter, that he had actually failed the test. The letter indicated he was to enroll in another school in the fall. All of the choices were adult education programs, awarding high school equivalency diplomas. Several of Danny's peers received similar letters but the exact number is unknown. Danny had always been a decent student, consistently receiving B's and C's. He had already had his senior year book photo taken and was planning to go to college. (p. 42).
24. National perspective	New York Times,Diana Jean Schemo, (January 18, 2004).	9th grade key to success, but reasons are debated	Article discusses research by Dr. Walt Haney that shows a dramatic increase in student attrition between the ninth and tenth grades over the last 30 years. Haney attributes this dramatic increase in the number of students who graduate from high school to two trends: increasing course requirements and growing demands that high school students pass specific standardized tests, otherwise known as exit exams, to receive a diploma.

This document is available on the Education Policy Studies Laboratory website at: **http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc**

Table 3, Continued

25. Florida	St. Petersburg Times, Editorial (June 18, 2004).	From FCAT to GED	Data shows more students, especially younger students, are opting for the GED instead of FCAT approved degrees. Asked to explain why 21,000 teenagers took the GED test last year, an increase of 78 percent in one year, Department of Education spokesman MacKay Jimeson pointed to state requirements for credit hours and grade point averages. But those standards haven't increased in almost a decade. So, Jimeson added: "Unfortunately, we still have kids in the system who have been affected negatively, by the old system, which did not have high standards and accountability." Last year, for example, nearly 1,000 of the students taking the GED were 16 years old. Given that the high school exit FCAT is first given in tenth grade, is it possible that early failure on the test is leading students to drop out of school? Debby VanderWoude, administrator of Dixie Hollins Adult Education Center, says she may be seeing such a trend in Pinellas. "I'm not surprised," she told a reporter. "We've been seeing a whole lot more of the younger ones."
26. San Antonio, Texas	San Antonio Express-News, Lucy Hood, (November 16, 2000), p. 1A.	Fewer Texans are finishing high school	Article about the decrease in the percentage of Texans who have completed high school. NCES reported that fewer than four of five Texans between the ages of 18 and 24 years old held a high school diploma or an alternative certificate in the period from 1997 to 1999. The article includes comments from Walt Haney whose research has found that Texas has one of the highest rates of completion of the GED of anyplace in the country. "According to Ruben Olivarez, former dropout czar for Dallas Schools and current (in 2000) superintendent in the San Antonio School District, the high dropout rates and larger numbers of students pursuing a GED can be attributed to a combination of things – including the high stakes associated with passing TAAS, as well as the increased rigor of passing Texas' academic courses.
27. Arizona	Indian Country Today, Wahinkpe Topa, Associate Professor in Education Leadership (March 10, 2004).	Wahinkpe Topa says " No more!" to laws that hurt our children	Writer vehemently argues against the No Child Left Behind act as especially damaging to Indian children across America. The pressure associated with high-stakes testing under No Child Left Behind causes "teachers to all but ignore art, music, critical thinking, creative autonomy, and social environmental justice. In many states, the tests themselves are flawed. The pressure on children and parents 'not to fail' is creating serious issues in self-esteem, a problem already serious in Indian country." The writer goes on to argue, "Labeling schools and children as 'low performing' is putting them at risk for takeover, forcing them to adopt a curriculum that is not culturally related, and allow Indian and non-Indian educators who have been propagandized to reproduce a system of thinking about the world that is harmful to everything does diminish culturally related activities."
28. Trenton, New Jersey	Christian Science Monitor, Stacey Vanek Smith, (October 21, 2003).	One exam, fewer ways to try again	Story laments the plight of 17-year-old John Lassiter who had dropped out of school during his sophomore year, but had recently re-enrolled in a nontraditional program in Trenton. Although armed with new found confidence from success at his new school, a statewide exam might stand between him and receipt of a high school diploma. Currently, New Jersey has an alternate route for students who fail the statewide test to receive a diploma, but that route may soon be closed to many students like John.

This document is available on the Education Policy Studies Laboratory website at: **http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc**

Table 3, Continued

29. Falmouth, Massachusetts	Cape Cod Times, K. C. Myers (Staff Writer), (January 14, 2004).	Dream denied: Tracey Newhart's future is once again in limbo because she didn't pass the MCAS	Story about the plight of a teenager with Down's syndrome who, after failing the MCAS and not receiving an official diploma certified by the state (only one certified by the school district), was denied entry into a local culinary college. The college stated that they needed "an official high school diploma" to let her in. The district endured state-level battles over the legitimacy of a locally-distributed diploma. In Tracey's case, the district superintendent even acknowledged that the school didn't have time to prepare Tracey for the MCAS. Still, there would be no leeway to give Tracey an official diploma, even though she worked hard to fulfill all other academic obligations and had her heart set on culinary school. Of course, if Tracey graduated prior to 2003, she would have had a diploma since 2003 was the first-year students had to pass the MCAS to get a diploma. In Massachusetts, Tracey is among 28 percent of special needs seniors who had not passed MCAS as of May 2003.
30. Louisiana	The Times-Picayune, Rob Nelson, (March 17, 2004).	Testing the limits: New federal policy will see special-education students taking the LEAP test, but critics decry the move as unfair	The story debates the merits of the requirement in No Child Left Behind that students <u>not</u> be tested out of grade level. In Louisiana, one student is facing the pressures of having to take the standardized LEAP test at grade level. In 1998, Breion Jones was in a car accident that claimed the lives of her brother and sister and left her paralyzed from the chest down. The head injuries led to mental disabilities and forced her to relearn basic math and how to hold a pencil. In spite of her cognitive challenges, this student will be forced to take the state mandated Louisiana Educational Assessment Program test. According to the article, "This is the first year that federal policy changes will force some special education students to take traditional standardized tests instead of alternative tests designed for disabled students. The revised policy restricts what is known as 'out-of-grade-level' testing, causing teachers to instruct students on the grade level appropriate for their ages, but not according to their mental disabilities." This policy is creating some stress among parents and students who have disabilities. Breion's mother is quoted as saying, "I think (school officials) had good intentions, but they overlooked my daughter. It's just going to be another school day for us. I'm not going to put any additional pressure on her, because it's not her test." The principal of the elementary school at which Breion is enrolled also finds the policy as unfair. She is quoted as saying, "It's unfair and it's mean. We are hurting the children we are supposed to be helping the most."
31. Florida	Star Banner (March 4, 2004).	The trials of the test	Story about one 11-year-old student in Ocala Florida who was held back twice in his school life, making him one of the older third graders of his school. Last year, K'von Brown was held back again (after being held back in first grade to catch up with his classmates) and made to repeat the third grade. Then, after a clerical error on his report card, did not receive the summer school help he was entitled to that was designed to help those who fail the test. His mother worries that K'von will always be behind. She notes, "You've got a motto: No Child Left Behind. He's [her son] supposed to be in the sixth grade." As a result, each night, K'von's mother helps him with math and reading and stresses the importance of those subjects in every day life. But, according to his mother, K'von has grown more discouraged – even if he has good grades, he can't pass the grade if he doesn't pass the test.

This document is available on the Education Policy Studies Laboratory website at: **http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc**

Table 3, Continued

32. Florida	St. Petersburg Times, Rebecca Catalanello (Times Staff Writer), (April 24 2004).	FCAT's English trips non-native speakers	In New Port Richey, about 35 to 40 students protested outside the Gulf High School Friday morning angered over Florida's high-stakes testing requirements for seniors. In Florida, seniors must pass the test to receive a standard issue high school diploma. Some of the signs included exclamations such as "FCAT Unfair!" Another one said "I failed the FCAT by one point." One senior, Leah Hernandez, has a 3.5 grade point average. She came to the country with her family five years ago from Mexico and said her English skills are the main reason she won't be getting her diploma next month. According to the article, Hernandez took the state's reading test six times but ended up six points short of the cutoff. One student complained, "For (non-native English speakers), if you're off by a couple of points, there has to be some leeway." One of the protesters wanted members of the Florida State Legislature and Jeb Bush to take the FCAT and give the public their scores. According to this student, "Maybe they can get a good idea that it is not as easy as it sounds."
33. South Carolina	Associated Press (October 6, 2002).	Black students more likely to flunk PACT	According to a report documented by the Associated Press, black students in South Carolina were more than twice as likely than white students to flunk the 2002 Palmetto Achievement Challenge Test (PACT). The newspaper analyzed the 2002 PACT scores of 532 public schools where at least 20 percent and no more than 80 percent of the students are minorities. In just 18 of those 532 schools did black students match or outperform white students. Nearly 40 percent of black students at Merriwether scored in the top two levels of PACT, while fewer than 10 percent failed the test. Statewide, 14 percent of black students scored at PACT's top levels and more than 45 percent failed.
34. New York	WNBC.com (Posted May 19, 2004), http://www.wnbc.com/education/	Protesters: Standardized tests tool of segregation	Parents and teachers are fighting the growing use of standardized tests in New York. About 50,000 petitions were presented to key legislators arguing that high school Regents exams "foster a segregation that was supposed to have been ended 50 years ago under the Brown v. Board of Education decision by the U.S. Supreme Court." The article goes on, "Extensive reliance on high stakes Regents exams has turned public schools into test-driven institutions that emphasize the most menial skills," said Jane Hirschmann of Time Out From Testing. "This narrow focus perpetuates the educational gap that Brown (v. Board of Education) was designed to fix ... High-stakes testing is a way we keep 'separate and unequal.'" The group said minority students – who attend mostly under-funded schools – often fare poorer on the standardized exams than their white counterparts because the tests are biased and graded on a curve that could fail them or prompt them to drop out.

This document is available on the Education Policy Studies Laboratory website at: **http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc**

Table 3, Continued

35. Massachusetts	Milford Daily News, Claudia Torrens (News Staff Writer), (April 29, 2004).	Lawsuit: MCAS unfair for minority and disabled students	Advocates for immigrants have filed a lawsuit claiming the MCAS is unfair to students with disabilities and to minority students whose native language is not English. "The Multicultural Education, Training and Advocacy Coalition, or META, along with other advocacy groups, claim that Limited English Proficient students should not be denied a high school diploma because of their level of English." The lawsuit was submitted on behalf of eight students who were denied a high school diploma because they did not pass the graduation exam. "Immigrant advocates also claim most minority students and students with disabilities attend schools identified by the state Department of Education as 'low' or 'underperforming,' which do not adequately prepare students for the test. … 'Minority students whose native language is not English are mostly concentrated in school districts that have been denying them a good education,' said David Godkin, partner of Testa, Hurwitz & Thibeault, LLP, one of the Boston law firms that filed the suit." In the suit, "Plaintiffs ask for students to be able to complete several areas of the MCAS in their native languages, not only the math section. They also want the appeals process to include all students, not only those who are disabled."
36. Alaska	Heidi Loranger from Anchorage 11 News, web site address http://www.ktva.com	Lawsuit filed against exit exams	According to the reporter, exit exams destroy children's futures. "That's one of the reasons why students with disabilities filed a Class Action lawsuit against the testing Tuesday. The lawsuit is specifically by students with disabilities, but it could have a far-reaching effect." In Alaska students must pass the exit exam in order to receive a regular diploma. The lawsuit claims that, "More than 500 students are not being given a fair chance to pass the test. In this suit, five students were named … however, they represent more than 500 disabled students who are qualified, but likely will not receive a diploma this spring because of the exams." Advocates of the lawsuit say 75 seniors who have not passed the test are students with various learning disabilities such as dyslexia, emotional disturbance, speech and language delays.
37. Colorado	The Denver Post, George Merritt, Staff writer (March 17, 2004).	Special kids, standardized tests: Parents, educators question use of same yardstick for every student	In Colorado several educators are complaining about the federal legislation that requires almost all children to be tested by the same standards. As the article notes, if schools fail to show progress, harsh penalties will follow. The writer notes, "But assessing special-needs children with standardized tests is unrealistic because, in most cases, they can't perform as well academically as typical students, many parents and school officials say." The article goes on to quote Principal Gary Hein of Euclid Middle School in Littleton, Colorado, "I have no problem with being held accountable … I welcome it. But, a test will never increase these kids' cognitive abilities. There's just a reality to that. We would rather teach these kids to be successful in life than try to make them do well on a test." In Colorado, roughly 74,000 special needs students must be tested along with typical students. However, these special needs children have posted the lowest scores on such exams, according to the State Department of Education. The article notes, "Everyone involved with special needs students agrees, the children should be included in as much typical curricula as possible. But, they disagree on whether testing is another part of inclusion or a wedge driving special needs programs away. Parents fear testing will eventually lead schools back to segregating special needs students. They test. They fail. And because failing can carry harsh penalties for the school, they are shunned."

This document is available on the Education Policy Studies Laboratory website at: **http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc**

Table 3, Continued

38. Tennessee	The Commercial Appeal, Ruma Banerji Kumar (October 6, 2003).	Set up for failure?--State tests pose daunting challenge for special education	Teachers express growing concern for their special education students, who are forced to take tests that are above their intellectual ability. In Sallie Rushing's elementary classroom, her students (ages 14 - 17) struggle with severe learning, emotional, and mental disabilities. However, as high school sophomores, juniors and seniors, they are forced to take Gateway Algebra I and 11th grade writing tests, even though most of them are at elementary school level work. Across the Memphis city school district, about 15,000 of 16,800 special education students took regular TCAP and Gateway tests because of the No Child Left Behind law. Statewide, across the board, schools that didn't make adequate progress often had low special education scores, said Julie McCargar, the state education department's director of federal programs. "This is a very controversial part of the law," she said. State officials said the law's aim is to ensure special education students are being adequately educated and regularly tested. But educators across the spectrum agree the mandate to test special education students like mainstream students has daunting implications for schools already struggling to get off state "target" lists and to avoid tough penalties. The No Child Left Behind law also requires all students, including special education students, to perform at grade level by 2013.
39. Texas	San Antonio Express-News, Jenny LaCoste-Caputo (October 3, 2004).	"Most fragile" can break school's academic ranking	NCLB requires that only 1 percent of a school's population can take an alternative version of the state exams used to make accountability decisions. Yet, there are growing concerns on the effects of this policy when a greater proportion of the school population is designated as special needs. For example, at one San Antonio middle school 15 percent of the school's 1,100 student enrollment is designated as special needs students. Still, almost all of these students are required to take the version of the test given to all students. As a result, growing numbers of schools are being labeled as failing because either (a) not enough students are being tested or (b) the special education students' scores are bringing down the school average because they are forced to take a test that is likely beyond their abilities.
40. Chicago	Chicago Tribune, Darnell Little (October 20, 2003).	Study maps city's teen dropouts without jobs	A study released by Chicago-based association of independent schools reports that nearly 16 percent of Chicago teens ages 16 to 19 are without jobs. Twenty percent of black and 18 percent of Latino youth are unemployed, compared with seven percent unemployment rate for white youth.

This document is available on the Education Policy Studies Laboratory website at: **http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc**

Table 3, Continued

41. Birmingham, Alabama	The Birmingham News (Editorial) (June 16, 2000).	High-stakes testing wobbly: Integrity not a reason to give up	One city school board member in Birmingham, Alabama alleged full-scale corruption on the SAT 9 because huge numbers of students were "withdrawn" from the roles in the months leading up to Alabama's Stanford 9 test. According to the story, a full 10 percent of Woodlawn High School's student body was yanked. According to the editorial, "coincidentally, or maybe not, Woodlawn is one step away from state takeover, the ultimate consequence for failure to improve test scores." The principal says these disruptive, uninterested, and unmotivated students would have been withdrawn even without the SAT 9, which followed closely the "withdrawal" of the students The editorial concludes with the question, "What are we to believe if, when the 2000 test scores are released later this month, Woodlawn makes a miraculous recovery from Alert 2 status to Academic Clear? We already know the very students who need school the most were banished from it."

This document is available on the Education Policy Studies Laboratory website at: **http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc**

Table 3 presents data on dropouts and push outs from our schools. The situation in New York City was so bad and so obvious that the Chancellor of the system embarrassedly admitted that for many years the NYC schools engaged in the systematic dumping of thousands of children onto the streets (see Article 2). Birmingham, Alabama seems to have been caught in a particularly egregious violation of youth rights so they could game the testing system (see stories in Article 23 and Article 42). Just before administration of the annual high-stakes tests, Birmingham officials had 522 young people "administratively withdrawn" from high school. By doing so scores on the state test went up and the district superintendent received a substantial bonus and pay raise while several schools avoided being taken over by the district.

Just before administration of the annual high-stakes tests, Birmingham officials had 522 young people "administratively withdrawn" from high school. By doing so scores on the state test went up and the district superintendent received a substantial bonus and pay raise while several schools avoided being taken over by the district.

There is another route that can be taken to get students off the official roles. That is by encouraging them to enter G. E. D. (General Education Degree) programs. K. W. Arenson, reporting in the New York Times[47] notes that this program is growing and removes weak students from school accountability programs:

> ….nationally teenagers accounted for 49 percent of those earning G.E.D.'s in 2002, up from 33 percent a decade earlier. The growth has been especially pronounced in New York City. Last year, more than 37,000 school-age students were in G.E.D. programs run by the school system, up from 25,500 two years earlier.

This document is available on the Education Policy Studies Laboratory website at:
http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc

> Experts attribute the flood of young people in part to.…the increased difficulty of earning a traditional high school diploma in many states. New York, for example, has made passing five Regents exams a condition of graduation, and no longer offers a lesser diploma for weaker students.
>
> Under the federal No Child Left Behind law and state efforts to hold schools more accountable, schools have more incentive to discourage weak students from staying. Students who transfer to G.E.D. programs are usually off school rolls, but in many states are not counted as dropouts.
>
> Mr. Chaplin, of the Urban Institute, said he had "found pretty strong evidence that the G.E.D. option has been encouraging kids to drop out of high schools nationwide. The rules governing the G.E.D. have become more lenient over time," he said. "Under No Child Left Behind, we're holding schools very strictly accountable for test scores, but barely holding them accountable for students who drop out or go into G.E.D. programs. It is like holding hospitals accountable for the condition of patients who leave, but ignoring the number who die. It's a perverse incentive system."

Article 27 in this table makes an important point about the curriculum imposed by high-stakes testing for Native Americans. Not only is the curriculum narrowed to focus on the tests, an issue we discuss below, but the curriculum is culturally irrelevant to the students it is supposed to educate. In an era in which our nation is trying to teach teachers to use culturally relevant pedagogy with Native Americans, African Americans and Latinos, teachers feel forced by high-stakes testing pressures into designing curricula that is boring and alien to many of their students. What our policy is doing, then, is

This document is available on the Education Policy Studies Laboratory website at:
http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc

driving out these students with a test-oriented curriculum and abandoning the kind of culturally relevant curriculum that might keep them in school. This will have the effect of driving up the already high dropout rates for students from other cultures. The flip side, of course, is the scores in the schools and districts go up when students leave, making it especially difficult to persuade school personnel to address this issue since it is to their benefit to drive these students out.

Articles 29 and 30 illustrate a different kind of problem: The need to test has replaced the need to care, a corruption of the traditional role of teachers. In these horrific cases students are not forced out of school, they are forced to stay in and endure a system that is cruel to them. In both these cases we see the problems that occur when laws are written by policy makers a long distance from the places where the laws must be administered. Under high-stakes testing programs these students are being hurt while teachers and administrators are made complicit in the acts of cruelty. The interpretation of the rules for testing seems particularly harsh and bureaucratic in Florida. The South Florida Sun Sentinel reports on the state's zero excuse policy for children, who *must* take the test in spite of frightful emotional trauma or face being held back a grade.[48] For example, no flexibility in test procedures was allowed for a 14-year-old child who lost her brother to a shooting and was still mourning; or for the 15-year-old who had previously found his father hanged in their home and suffered anxiety attacks from that event. In fear that students and their parents would take advantage of any leniency in the rules, the school system and all the personnel in it become oppressors of some of the children.

This document is available on the Education Policy Studies Laboratory website at:
http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc

It appears to us that the most important problem emerging from reading these stories is the loss of humanity that high-stakes testing produces. Not just in the treatment of the special education students and those with special psychological problems, but also in the treatment of poor, immigrant, and minority children who are likely to be score suppressors not score increasers. Teachers and administrators are beginning to treat the score suppressors and the score increasers in the same way that children on the playground treat them. When forming teams in baseball, basketball, and track, the score increasers (the best hitters, shooters, and runners) are the ones who are picked first. Those perceived as the score suppressors are picked last, or are not picked at all. They endure a public humiliation. Now the same incentive system is found inside the classroom. Some schools and districts are treating the academically weak among their students as pariahs. They are to be discarded or labeled in some way so as not to take the tests and bring down the average.

In some schools that are trying to make Adequate Yearly Progress, these students get all the schools' attention, including special tutoring. The students who are performing less well academically are score suppressors—they get no resources.

This problem was made obvious in a story carried by The Los Angeles Times.[49] It discusses the "cusp" kids—those almost at the point of passing the high-stakes test, perhaps needing a little extra teaching time to help them to pass. In some schools that are trying to make Adequate Yearly Progress, these students get all the schools' attention, including special tutoring. The students who are performing less well academically are score suppressors—they get no resources. The schools have effectively written them off.

This document is available on the Education Policy Studies Laboratory website at:
http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc

The story makes clear that the increased importance of achievement indictors has taken away some of the humanity of these educators. This can also be seen in some of the stories in Table 3, such as those summarized in Articles 37 and 40.

In a set of stories we found about the gifted we discovered how they too are ignored or treated simply as score increasers, rather than sentient students with their own ambitions. For example, in Arizona the gifted in some schools are written off because their scores, like those of the least able, are so predictable. It is the cusp students who get the schools' attention. Thus in high schools that must confront high-stakes exit exams, special instruction for the gifted is often ignored.[50] But in Arizona, perhaps elsewhere, the gifted are not totally ignored. Because they are score increasers they sometimes are ill-treated. Schools and districts in need of higher scores talk their most talented students into taking the high-stakes high school exit exam over and over again, after they pass it![51] This, of course raises school and district scores, though it wastes a students' time. It would appear that students are less important than their test scores.

Still another problem concerning the score increasers, the academically talented in a school, was revealed by The Wall Street Journal.[52] They reported on an Ohio sixth grader who was in a gifted program but his scores on the state's high-stakes test were never counted in the school he attended, a school for the gifted. Instead, his scores were credited to the neighborhood school he does not attend so the average score of that school could go up. If no "credit" was given to the local schools, the local schools might not identify any students as gifted, fearing that they might lose the student to a school with a gifted program. Apparently such "credit" systems exist in Missouri and Iowa, where their schools also fear losing their high scorers to another school. The scores of the gifted

This document is available on the Education Policy Studies Laboratory website at:
http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc

become something to negotiate over, and if credit were not given such students might never be identified as gifted at all! As we saw with the least able students, the score, not the child, has become more important. This is a serious distortion of the roles our nation wants our schools and teachers to enact.

Misrepresentation of Dropout Data

We noted, above, that dropout data was itself an indictor of a district's productivity. So when a district increases its dropouts and pushouts to influence its test scores, it then has the problem of presenting its dropout rate for public scrutiny. As Campbell's law might predict, some districts will lie about those data. Articles 3 and 4 of Table 4 show how Houston, in particular, misrepresented data on dropouts. In Houston, a Sharpstown High School dropout and his mother noticed that the high school he should have been attending had no dropouts recorded for that year.[53] Since that was obviously not true, the mother, with the help of a local newspaper, began an investigation. She found that 462 students had left Sharpstown and all were reported to be in charter schools or other schools, though Sharpstown administrators had not asked these students where they were going and had no knowledge of what they were doing. None were recorded as having dropped out! Sharpstown had started with about 1,000 freshman and its senior class was about 300, but had recorded no dropouts. In 2000-2001, the year that Houston said it had a 1.5 percent dropout rate, about 5,500 left school and over half

Houston's Sharpstown High School had started with about 1,000 freshman and its senior class was about 300, but had recorded no dropouts.

This document is available on the Education Policy Studies Laboratory website at:
http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc

should have been counted as dropouts, but were not. For his leadership of the Houston school district, Superintendent Rod Paige was honored by McGraw Hill, and on the basis of his record as a school leader, he was elevated to the highest educational position in the land, Secretary of Education under President G. W. Bush. For their outstanding urban education programs Houston received $1,000,000 from the Broad Foundation. And previous to the announcement that it was not telling the truth, Sharpstown high school staff received bonuses based on good attendance, low dropout rates, *and increased test scores*. Shapstown and Houston are models of the applicability of Campbell's law.

Houston also had another indictor of success that it was quite proud of—the rate of college attendance by its high school graduates. D. J. Schemo of the New York Times reports:[54]

> At Jack Yates High School here, students had to make do without a school library for more than a year. A principal replaced dozens of experienced teachers with substitutes and uncertified teachers, who cost less.
>
> And yet from 1998 to 2002, Yates High reported that 99 percent to 100 percent of its graduates planned to attend college.
>
> Across town, Davis High School, where students averaged a combined SAT score of 791 out of a possible 1600 in 1998, reported that every last one of its graduates that year planned to go to college.
>
> Sharpstown High School, a high poverty school that falsely claimed zero dropouts in 2002, also reported in 2001 that 98.4 percent of its graduates expected to attend college.

This document is available on the Education Policy Studies Laboratory website at:
http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc

These reports were completely false. As with the dropout scandal, Secretary of Education Rod Paige refused to comment. One of Houston's principals, who asked not to be identified for fear of reprisals, said this:

> Lower-level administrators inflated their figures in the hope of attracting the children of active, involved parents. More students also mean more money from the state. On paper, her school claimed that almost all of its graduates were headed for college. In fact, the principal said, most of them "couldn't spell college, let alone attend." [55]

It might not be merely coincidental that Enron was headquartered in Houston. But Houston is not alone, as Table 4 makes clear.

This document is available on the Education Policy Studies Laboratory website at:
http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc

Table 4: Misrepresentation of Dropout Data

Location of Story	Source	Headline	Story
1. New York	New York Times, Tamar Lewin & Jennifer Medina (July 31, 2003).	To cut failure rate, schools shed students	Many of New York's struggling students were pushed out of the school system and classified under categories that hide their failure to graduate. According to the article, "Officially the city's dropout rate hovers around 20 percent. But critics say that if the students who are pushed out were included, that number could be 25 to 30 percent. The city data make it impossible to determine just how many students are being pushed out, where they are going and what becomes of them. But experts who have examined the statistics and administrators of high school equivalency programs say that the number of "pushouts" seems to be growing, with students shunted out at ever-younger ages. Those students represent the unintended consequence of the effort to hold schools accountable for raising standards: As students are being spurred to new levels of academic achievement and required to pass stringent Regents exams to get their high school diplomas, many schools are trying to get rid of those who may tarnish the schools' statistics by failing to graduate on time. Even though state law gives students the right to stay in high school until they are 21, many students are being counseled, or even forced, to leave long before then. And yesterday, after declining to comment on the issue for two months, Chancellor Joel I. Klein conceded the point. "The problem of what's happening to the students is a tragedy," he said, "It's not just a few instances, it's a real issue."
2. Texas	Education Week, Jeff Archer (September 2 4, 2003).	Houston case offers lessons on dropouts	An analysis sponsored by Ed Week found that about half of the 108 schools throughout Texas where 70 percent or more of students were considered at risk of academic failure claimed a dropout rate of one percent or less. However, a closer examination revealed that many schools' enrollment actually decreased by 30 percent or more by the time ninth graders reached 12th grade.
3.Houston, Texas	New York Times, Diana Jean Schemo (July 11, 2003).	Questions on data cloud luster of Houston schools	Story about continuing dropout debacle of Houston. A recent state audit found that more than half of the 5,5000 students who left during the 2000-2001 school year from 16 middle and high schools should have been declared as dropouts, but were not. One auditor noted that "many students who had left school were coded as intending to enroll in an alternative or high school equivalency program, and were, by Texas rules, not dropouts." This coding, however, was mostly based on principal reports.
4. Houston, Texas	Houston Chronicle, Staff Writer (July 13, 2003), p. 2.	To hold accountable: Get rid of cheaters, not education accountability system	Article talking about how many high schools are increasing the number of students who are exempted from the test, but assigning the reasons for the exemption as "other." At one school, where administration is already under investigation for faking low 2001 dropout data, "23 percent of the schools' tenth graders did not take the TAAS during the '00-'01 school year. Nearly four percent (five times the district average) were absent and more than four percent were exempted because of limited English or because they were special education students. A whopping 15 percent missed the test for 'other' reasons."

This document is available on the Education Policy Studies Laboratory website at: http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc

Table 4, Continued

5. Pleasanton, California	Tri-Valley Herald, Jill Tucker (April 27, 2003).	State issues annual data on dropouts	California's dropout rates are misleading. High schools are responsible for self-reporting dropout figures which means that schools vary widely in how they define "dropout." Dropout rates are estimated to be around 11 percent; however, graduation rates suggest that 32 percent of freshman who started high school in 1998 didn't graduate last spring.
6. New York	Newsday, John Hildebrand (December 19, 2003), p. A26.	Freeport's dropout problem; State audit: Rate higher than district said	Student dropout rates at Freeport High School have run more than five times higher than reported by the district, according to a recent state audit that traces the miscount to the high school, as well as an alternative evening school for troubled teens. The state probe, completed in November as part of an annual audit of districts, found that as many as 255 students or 10.4 percent of the high school's enrollment dropped out during the '00-'01 school year. In contrast, district officials had reported to the state losing only 46 students for the '00-'01 year. The audit by the State Education Department found that 175 students, not 46, should have been counted as dropouts.
7. Massachusetts	Boston Herald, Kevin Rothstein (August 23, 2003).	Education expert testifies MCAS pass rate misleading	Walt Haney presents data to show that state wide results on MCAS is misleading because a disproportionate number of low-income students did not take the test. In a poor district, just 48 percent of freshman class of 2003 made it to their senior year. In contrast, 97 percent of students in a wealthier district made it.
8. Massachusetts	Metro West Daily News, Peter Reuell (Staff) (June 13, 2004).	Test numbers draw criticism: Some wonder if latest DOE stats paint false picture	For accuracy's sake, the state's reports ought to cover more than just who passed. "(They should say) we're going to report what the performance of the state is broken down by the percent who pass MCAS, the percent who don't pass MCAS, the percent who drop out, the percent who are repeating a grade, the percent who move to another state and the percent who transfer to another school."

This document is available on the Education Policy Studies Laboratory website at: **http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc**

Misrepresenting data occurs also for other indicators. In Massachusetts, there is strong (and vocal) opposition to the high-stakes testing of student performance on the Massachusetts Comprehensive Assessment System (MCAS). Therefore, it was no surprise that when the state released data on all those passing the test it was examined closely. It was certainly in the state's interest to have those data reassure the public that the testing system was working.

Ann Wheelock and colleagues of the National Board on Educational Testing and Public Policy at Boston College noted that the state claimed that 96 percent of the class of 2004 had passed the MCAS and thus would graduate high school.[56] But what the state had failed to do was disclose the dropout rate.[57] The class of 2004 had lost 18,232 students from grade 9 to grade 12. That represents a 23 percent dropout rate and makes the class of 2004's actual passing rate 74 percent, not 96 percent. Moreover, if one considers the cohort beginning at grade 9, instead of just the survivors at grade 12, then the pass rate for African Americans and Latinos was 59 percent and 54 percent, respectively. This way of reporting the data paints a different picture.

The state of Massachusetts did the same thing when reporting the numbers of special education students that passed the state tests. Its press release reports that 5,268 out of 7,307 grade 12 special education students had passed the MCAS and would graduate that spring.[58] The state declared the passing rate of these students to be 72 percent. But once again the state left out the non-survivors, the many pushouts and dropouts that left along the way. Using ninth grade special education enrollments the pass rate was recalculated by Ann Wheelock and colleagues as 42 percent, not 72

This document is available on the Education Policy Studies Laboratory website at:
http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc

percent.[59] Less than half of the class of 2003 survived and graduated, thus inflating the MCAS pass rate substantially.

As Campbell's law predicts, indicators that take on inordinate value are the most likely to be corrupted. It appears that the Massachusetts State Department of Education was not presenting all the data to the public. They did not engage in a conspiracy to deceive the public, as appears to be the case with the administrators in Houston. But they nevertheless did not tell the entire story, telling only the piece of the story that made them look good on the indictor of interest.

Teaching to the Test

It is true that if we have a good test, then teaching to that test is appropriate. Instruction and assessment would be aligned, as they should be. But there is a fine line between teaching to the test and teaching the test itself. That line is often hard to see and it is possible that some educators and test preparation companies deliberately cross that line. This would corrupt the indictor, making it quite difficult to interpret any assessment so compromised. Table 5 presents some stories concerned with this issue.

This document is available on the Education Policy Studies Laboratory website at:
http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc

Table 5: Teaching to the Test

Location of Story	Source	Headline	Story
1. California	Pasadena Star-News, Marie Leech (October 13, 2002).	Teachers question testing frenzy	Teachers complain about the state's rigorous testing schedule that forces them to spend more time preparing students for the test rather than teaching in ways they would prefer. According to the article, "In Larry Wiener's fifth-grade class, students don't get to do a lot of drawing and painting. They don't spend much time learning about science and social studies. But they do know how to put points on a graph, something they will see again this spring on the CAT-6 achievement and California Standards tests. "Wiener, like many other teachers, say the state's rigorous testing schedule is failing students because it leaves teachers no choice but to base instruction on what's being asked on the tests. Many experts agree, saying the state puts so much emphasis on standardized test scores that teachers often spend huge amounts of time on math and English – two of the main subjects tested – allowing other subjects like the arts and sciences and social studies to fall by the wayside. "Tests have become the holy grail of education, said Wiener, a teacher at Ynez School in Alhambra. Because he has to focus so much on math, reading and language arts, Wiener has had to 'downplay many other subjects, especially the arts.'"
2. Texas	New York Times, Diana Jean Schemo & Ford Fessenden (December 3, 2003).	Gains in Houston Schools: How real are they?	The article highlights the plight of one teenager who was trained to write five-paragraph "persuasive essays" for the state exam, but was stumped by her first writing assignment in college. She failed the college entrance exam in math twice, even with a year of remedial algebra. Are the successes in Texas real or imagined? The article argues that in national comparisons on parallel tests, Texas is not gaining ground academically and this story illustrates concerns that the accountability system is forcing a narrowing of the curriculum and a lack of transfer of information.
3. Manchester, New Hampshire	Union Leader Damian J. Troise, (June 23, 2003), P. B1.	'No Child' rules bring problems	An article about a memo sent to the school board of one local district from the Curriculum Advisory Committee in which worries were expressed that the increased time teachers would spend on preparing students for a new annual test in math and language arts might pressure teachers to "teach to the test." Members worry that students are spending too much time on tests. In the state, the average sixth and tenth grader spends up to 25 hours each year preparing for state exams whereas third graders spend 18. The committee worries also that the high stakes attached to the exam may tempt some teachers to "cheat" or at least engage in behaviors that will sway the results. Another worry is that many of their finest teachers would be considered not "qualified" by NCLB standards.

This document is available on the Education Policy Studies Laboratory website at: http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc

Table 5, continued

4. West Virginia	Charleston Gazette, Linda Blackford (Staff writer), (July 30, 1005).	The testing dilemma: Should students be coached for a standardized test?	In West Virginia, the state labels schools as performing above or below academic standards based on how students do on a standardized test called the Comprehensive Test of Basic Skills. As a result, schools can be put on probation or encounter state monitoring and other forms of sanctions if they continue to yield low-performing test scores. As a result, schools have been forced to engage in practices that align the curriculum more strongly with the test and develop students' test-taking skills – some might argue this practice has led to increased teaching to the test. According to one county assistant superintendent whose district was under state probation, he has been forced to change the district level curriculum to help students' test scores. He notes, "The state established the process, and we can't opt out of the process … we're going to do some curriculum alignment but we're not going to reduce hands-on activities for the kids." He also had to make a list for teachers to follow as a strategy for helping to improve student test scores. Included in this list are: 1) Practice answering problems on a worksheet using a grid. 2) Know what is tested on CTBS and teach these concepts early in the year and then review all year. 3) Use test preparation materials throughout the year, a little bit each day. Talk about why other choices are not right. 4) Practice testing comprehension using multiple-choice formats. According to the writer of the article, there is a fine line between preparation and cheating and it is sometimes hard for educators to distinguish. For example: "Copies of test booklets have made the rounds of county school systems, teachers say, and after five years, they have a pretty clear picture of the test."
5. West Virginia	Charleston Daily Mail, Rebecca Catalanello, (September 9, 1999).	Educators debate role of testing	During the spring, in West Virginia, students are engaged in a week long standardized testing assessment. Many teachers throughout the state of West Virginia have expressed the concern that the effects of testing in the state and the pressures associated with the testing have made an indelible impression on their day-to-day instructional decisions. According to a retired assistant state superintendent of schools, Keith Smith, "The test shouldn't drive everything that's going on in the classroom and it doesn't … but you have to be reasonable. If you look at the test data, those are things you want people to know." In 1999 the State Department of Education introduced a set of curriculum standards that were to enter each classroom's instructional design. These standards were designed to parallel the subjects tested on the Stanford 9 standardized achievement test that is given to students in grades 3 through 12 and scores of which are used to make judgments about how good schools are doing in educating its students. However, as a result of the pressures, many teachers suggest that it is changing how they teach. For example, one sixth-grade history teacher said that before standardized testing took center stage she would arrange her sixth-grade history curriculum to explore world history. However, now she makes sure the students receive a unit on American history – a subject that students have just studied in the fifth grade – just because it will be on the test. She notes, "Quite frankly, I concentrate on the things that are going to be tested. I know if I'm going to survive, that's what you do."

This document is available on the Education Policy Studies Laboratory website at: **http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc**

Table 5, continued

6. Mississippi	The Clarion-Ledger, Cathy Hayden (April 30, 2004).	Much, including jobs, riding on public schools' annual tests	Schools across the state of Mississippi, as they gear up for the statewide Mississippi Curriculum Test, are spending a lot of time preparing for the test. And, the stakes are high in Mississippi. "If students don't do well, the poor performance reflects on the school, including administrators and teachers. Eventually, they could lose their jobs." And according to Susan Rucker, associate state superintendent of education, "You can't help but expect people to be a little tense and uptight about it. There are a lot of high stakes about it for everybody involved." The writer notes, "For Brinkley and other lower-performing schools, the pressure is excruciating. … Test preparation mania is playing out across the state, in many schools."
7. West Virginia	Charleston Gazette, Linda Blackford (Staff writer), (May 23, 1996).	Coaching improves test scores	Article discussing the efforts the state has taken to increase students' test scores. Although considered unethical and unhelpful to coach students on the statewide test, the Comprehensive Test of Basic Skills, repeated criticisms for low scores prompted the administration to hand out copies of "Scoring high" a workbook with practice tests for students. William Mullett, the county's director of counseling gave the credit for improved scores to teachers' hard work and narrowing the curriculum with the help of "Scoring High." "We resisted that effort for a number of years," Mullett said. "Finally, because of what other school systems were doing, we felt our students were being compared unfavorably. We leveled the playing field a little bit."
8. Texas	Houston Press, Shaila Dewan, (February 25, 1999).	The fix is in: Are educators cheating on TAAS? Is anyone going to stop them?	Story about allegations of cheating on Texas' statewide TAAS exam. Story uncovers allegations of cheating and "helping" behavior aimed at increasing student performance on tests. One school, in an effort to increase student test performance, focused on "using test-taking 'strategies' such as underlining key words and drawing arrows when rounding numbers. Children who used the strategies on practice tests were rewarded with, among other things, movies and popcorn." But, that wasn't all they got, according to another teacher who was instructed to "give out candy during the test to children who were using the strategies. If the child wasn't using them, the teacher whispered to the child 'use the strategies.'"
9. Boston, Massachusetts	Boston Globe, Kim Marshall (June 28, 2003), p. A1.1.	The junk food of education	The author argues that too many educators may fall victim to test prep strategies in an effort to provide a quick fix solution to student low achievement on the statewide MCAS. She worries that test prep materials are necessary in today's high stakes climate, but argues that educators must be careful on what kind of test preparation they provide their students. Too much drill and kill or teaching to the test via old MCAS items will only be harmful to students and the meaningfulness of test scores.
10. Austin, Texas	Austin American-Statesman, Jason Embry (Staff) (December 23, 2003), p. B1.	Teachers' petition aims to find her accuser	A former elementary school teacher filed court papers in an attempt to force the school district to name the parent who complained she illegally coached students on a standardized test. The teacher ended a 35-year tenure with the district after a parent accused her of helping a fourth-grade student on the TAKS. She was forced to resign, but now works at another district. The teacher was accused of encouraging students to sound out words and read over test questions, but the teacher denied providing any answers. Teachers who are found to have helped students on the state test could lose their certification.

This document is available on the Education Policy Studies Laboratory website at: **http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc**

Table 5, continued

11. Rochester, New York	Katrina Irwin, (January 26, 2004) (WROC-TV8), online: http://www.wroctv.com/news/story.asp?id=12098&r=l	High school students get another shot at exam	Last year, more than two-thirds of graduating seniors failed the math portion of the Regents exam. According to the news report, "That's when school administrators started asking questions. Rallies ensued and an independent panel looked over the test. They decided it was unfair and the results were thrown out. This year's version was supposed to be more focused so teachers would know how to prepare students. Shaina Colenan says this year's test was better. 'Some of the stuff last year, I don't think they taught us in school, but this year they taught us more of what would be on the test.'"
12.Indiana	Associated Press (June 19, 2002).	Retired math teach admits to giving students practice ISTEP tests	A Roosevelt math teacher admitted to investigators that she gave students practice tests with real ISTEP questions on it. Clark allegedly said nine different practice tests had been in circulation for years. She denied during an interview with Gilmartin that she had a copy of the spring ISTEP tests, which mirrored many of her practice questions. In one scenario, all the questions for the applied math skills section had been given to students ahead of time. Some secure multiple choice questions were also circulated and found their way into the hands of investigators. Clark said she gave out copies of questions during a crunch math session a few days before ISTEP.
13.Houston, Texas	Houston Chronicle, Annette Baird (Staff Writer) (April 27, 2004).	Spring Branch ISD forum offers insight into platforms	Candidates who were vying for the Spring Branch school district's board of trustees were questioned regarding their position on a variety of topics including school safety and accountability. "On the question of testing and accountability, candidates said they were proponents of accountability, but some that they are concerned about the amount of time pupils spend testing or preparing for the test. Both Haglund and Converse said the key is to make sure the right curriculum is in place. Haglund said the district is in the process of aligning the curriculum with the state tests and also going beyond that. Mandell said he is concerned that the district is too focused on teaching the test."
14.Dallas, Texas	Dallas Morning News, LaKisha Ladson (February 19, 2004).	Year of preparation faces test next week	Story about how Rockwall County school district officials have attempted to streamline the curriculum to aide teachers in curriculum decisions that are better aligned with the TAKS. In Texas, the Texas Essential Knowledge and Skills (TEKS) represents the curriculum objectives that are to be met in order to do well on the Texas Assessment of Knowledge and Skills (TAKS). Royse City district developed and implemented an "elaborate way to ensure that teachers are focusing on material that will be tested." District officials developed worksheets for the teachers that list each state objective and when it should be taught. They also provide room to spell out specific teaching strategies.

This document is available on the Education Policy Studies Laboratory website at: **http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc**

Table 5, continued

15. New York	New York Newsday, Wil Cruz (Staff Writer) (April 28, 2004).	1,800 will retake promo exams	Chancellor Joel Klein says that about 1,800, 8- and 9-year olds are in line to retake the English test. According to the report, "a department investigation found that 1,300 students who took the exam April 20, when it was originally given, were prepared with questions from last year's test." For third graders in New York, their performance on math and reading exams determine whether they are promoted to the next grade. For students who don't retake the exam, "their test results – and consequently, their promotion – will be determined by 30 of the 50 questions on the test.

This document is available on the Education Policy Studies Laboratory website at: **http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc**

Many articles make it clear that too much time is given over to test preparation. This has become a big business. Test preparation companies provide tutoring for individual students as well as offering large, formal curricula packages for test preparation to schools and districts. These private enterprises are driven by profit and that means their test preparation strategies need to "work." To work well the test preparation company needs items and an understanding of the test format that are very close to the actual items and formats used on the tests themselves. Thus there are incentives to cross the line and prepare students too directly for the test.

Test preparation companies provide tutoring for individual students as well as offering large, formal curricula packages for test preparation to schools and districts. These private enterprises are driven by profit and that means their test preparation strategies need to "work."

For example, in Tacoma in 1995, CTBS scores were at the 42nd percentile and rose to the 63rd in a few months. Superintendent Rudy Crew was then hailed as a miracle worker, received a large bonus, and on the basis of this work was then promoted to the Chancellorship of the New York City Public Schools. But to help get scores up Dr. Crew had hired a "test prep" consulting firm and the students were drilled on tests very much like those on the CTBS. Many practice tests with suspiciously similar items to the CTBS were used. The Tacoma miracle stopped (i.e. scores dropped) after Dr. Crew left Tacoma and the test prep company stopped working there.[60]

Both the time allotted and the trivial nature of test preparation programs irks teachers. To supplement the table we cite research articles about these issues, beginning with documentation from Texas where they were teaching a single writing format called,

This document is available on the Education Policy Studies Laboratory website at:
http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc

"the five paragraph persuasive essay." In this approach students are taught that each paragraph has exactly five sentences: a topic sentence, three supporting sentences and a concluding sentence much like the introductory sentence. The teachers call this "TAAS writing," as opposed to "real writing." Linda McNeil and Angela Valenzuela studied this issue and say:

> Teachers of writing who work with their students on developing ideas, on finding their voice as writers and on organizing papers in ways appropriate to both the ideas and the papers' intended audience, find themselves in conflict with this prescriptive format. The format subordinates ideas to form, sets a single form out as '*the essay*' and produces, predictably, rote writing. Writing as it relates to thinking, to language development and fluency, to understanding one's audience, to enriching one's vocabulary, and to developing ideas, has been replaced by TAAS writing to this format.[61]

Although rubrics make sense intuitively, they raise validity issues. Rubrics standardize scoring, increasing the reliability that is so important for large-scale assessments. But the rubrics also standardize the writing that is to be scored. Standardized writing, however, is not necessarily good writing, because good writing features individual expression, which is not standardized. The standardization of any skill that is fundamentally an individual, unique, or idiosyncratic skill, complicates its assessment. This, of course, presents a validity problem because the assessment cannot then produce scores that support valid inferences about the writing achievement of our students. This is a basic construct validity problem. Under rubric writing and rubric

This document is available on the Education Policy Studies Laboratory website at:
http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc

scoring, our tests can end up measuring not the construct of *writing achievement*, but the construct of *compliance to the rubric*, which is certainly not the same thing![62]

McNeil and Valenzuela provide another example of corrupting the indicator and the educators in the system by having teachers narrow the curriculum through drill activities, so that their students will perform well on Texas' high-stakes tests. They say:

> … high school teachers report that although practice tests and classroom drills have raised the rate of passing for the reading section of the TAAS at their school (that's the Texas Assessment of Academic Skills), many of their students are unable to use those same skills for actual reading. These students are passing the TAAS reading section by being able to select among answers given. But they are not able to read assignments, to make meaning of literature, to complete reading assignments outside of class, or to connect reading assignments to other parts of the course such as discussion and writing.
>
> Middle school teachers report that the TAAS emphasis on reading short passages, then selecting answers to questions based on those short passages, has made it very difficult for students to handle a sustained reading assignment. After children spend several years in classes where "*reading*" assignments were increasingly TAAS practice materials, the middle school teachers in more that one district reported that [students] were unable to read a novel even two years below grade level.[63]

On this issue, the voices of teachers and researchers have been well documented by Jones, Jones, and Hargrove in their excellent book: *The unintended consequences of high-stakes testing*.[64] Borrowing from their work we quote Alfie Kohn who noted:

This document is available on the Education Policy Studies Laboratory website at:
http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc

> To be sure, many city schools that serve low-income children of color were second rate to begin with. Now, however, some of these schools in Chicago, Houston, Baltimore, and elsewhere, are arguably becoming third rate as the pressures of high stakes testing lead to a more systematic use of low-level, drill-and-skill teaching, often in the context of packaged programs purchased by school districts."[65]

A principal in another study reported:

> The accountability system has an impact on everything we do. To focus on specific basic skills, you have to drill. We would like to get away from drill and pounding stuff into kid's heads; they don't remember it the next year. But if the accountability looks at scores to judge school effectiveness, you can't take your eyes off of basic skills.[66]

Other researchers reported:

> Everywhere we turned, we heard stories of teachers who were being told, in the name of "raising standards," that they could no longer teach reading using the best children's literature but instead must fill their classrooms and their days with worksheets, exercises, and drills.[67]

Teachers had to make time for all of this test preparation and most were not happy about it. Eighty percent of North Carolina's elementary teachers reported they spent more than 20 percent of their total teaching time practicing for high-stakes tests.[68] Worse yet, almost one-third of those teachers reported spending more than 60 percent of their time practicing for the state's tests.

This document is available on the Education Policy Studies Laboratory website at:
http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc

Mary Lee Smith found that as testing approached in Arizona, teachers had to give up valuable teaching activities to increase time for test preparation.[69] One teacher of creative writing dropped that subject from her curriculum mid-year as she and her school turned their attention to getting ready for the test. Smith found that 32 percent of the teachers she studied said they were "required" to do test preparation, and 28 percent of these Arizona teachers said they begin their test preparation at least two months before the test administration. Similarly, one principal told his teachers not to introduce anything new in the six weeks before the exam.[70] He demanded review of material in the same format of the test.

Teachers in Florida also discuss teaching to the test. Their voices have been recorded:[71]

> Teacher A: "I can say one thing, if my kids learn one thing in third grade, it is this: how to pass a standardized test even if you are not familiar with the material. Now is that what our goal is? Perhaps we should revisit it."
>
> Teacher B: "I have seen that schools are teaching to the test (how can you not?) and that is not a true reflection of student abilities. This is only a reflection of the abilities of each school to teach effective test-taking strategies, not academics."
>
> Teacher C: "Schools aren't improving their academics as students score better on the FCAT. They are just taking more time to teach to the test and unfortunately, away from real learning. We aren't getting smarter students, we are getting smarter test takers. That is NOT what we are here for!...The schools

This document is available on the Education Policy Studies Laboratory website at:
http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc

who score well are focusing on teaching to the test at a very high cost to their students."

Teachers in Colorado said some of the same things.[72] They reported: using formats similar to CSAP in regular instruction to help familiarize students to the official format of CSAP; using commercially produced test preparation materials similar to CSAP and test items released by the state to prepare students for the CSAP; and spending time teaching test-taking strategies to prepare students for CSAP, with the teachers in low performing schools spending about double the time that teachers in high performing schools spend on test preparation. They were not at all happy about taking time away from what they knew to be a richer curriculum.

Harold Wenglinski's research provides a sadder note about the ubiquity of test preparation programs.[73] Wenglinski looked at 13,000 students who had taken the fourth grade NAEP tests in 2000. He found that frequent testing actually reduced scores on NAEP, and that emphasizing facts (over reasoning and communication) also reduced students' scores. Since these two characteristics are prevalent in most of the test preparation programs we have examined, it is quite likely that many of the activities engaged in during test preparation are counterproductive.

Related to the issue of teaching to the test directly is the issue of narrowing the curriculum so as to concentrate attention on just those subjects that are tested. The corruption of professional judgment about curriculum offerings due to the pressures of high-stakes testing is inappropriate and might be unhealthy for students. We address that issue next.

This document is available on the Education Policy Studies Laboratory website at:
http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc

Narrowing the Curriculum

Table 6 presents stories that are concerned with the tendency of schools and districts to reduce what they teach that is not on the test, or to remove from the curriculum subjects that are not likely to be contributors to test score growth. Curriculum activities, like children, can be considered score suppressors or score increasers and are judged in a purely instrumental way. For example:

> Nap time is a daily ritual for the pre-kindergarten students at countless schools across the country. But in the increasingly urgent world of public education is it a luxury that 4-year-olds no longer can afford? "Nap time needs to go away," Prince George's County schools chief Andre J. Hornsby said during a recent meeting with Maryland legislators. "We need to get rid of all the baby school stuff they used to do." [74]

The Waltham, MA Daily News reports that recess for elementary students has been cut one day a week to make time for the high-stakes testing and to be sure that they look good on NCLB.[75] The report says that some nearby districts have done away with recess at the elementary grades completely. The educators around St. Louis have cut back on recess and physical education, too.[76] When the schools in the city of Tacoma, WA, reiterated their ban on recess, the Seattle Post-Intelligencer Editorial Board was prompted to advise the school administrators to get a life![77]

A teacher at a district nearby Waltham expressed concern that lunch had been reduced at her elementary school to less than 15 minutes on many days so that more time could be spent on the rigorous curriculum areas, meaning the areas that are tested. The

This document is available on the Education Policy Studies Laboratory website at:
http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc

school had started serving finger food—wraps and chicken nuggets—to get the students in and out of the cafeteria faster![78]

The Council on Basic Education is looking at bigger issues like the atrophy of the liberal arts in contemporary America.[79] Most of the areas they are concerned about (history, social studies, civics, geography, art and music, and foreign language) are not usually the focus of high-stakes testing. Therefore, under the pressure to succeed on the tests, these courses are abandoned. The loss of such courses was found to be greatest in minority communities.

History, social studies, civics, geography, art and music, and foreign language are not usually the focus of high-stakes testing. Therefore, under the pressure to succeed on the tests, these courses are abandoned.

The voices of teachers about this phenomenon are clear. It happens every time high-stakes testing is instituted. From Colorado we hear them say the following:[80]

> Teacher A: "We only teach to the test even at second grade, and have stopped teaching science and social studies. We don't have assemblies, take few field trips, or have musical productions at grade levels. We even hesitate to ever show a video. Our second graders have no recess except for 20 minutes at lunch."
>
> Teacher B: "I eliminated a lot of my social studies and science. I eliminated Colorado History. What else? Electricity. Most of that because it's more stressed that the kids know the reading and the math, so, it was pretty much said, you know, you do what you gotta do."
>
> Teacher C: "Those things (science and social studies) just fall to the backburner and quite frankly, I just marked report cards for the third grading

This document is available on the Education Policy Studies Laboratory website at:
http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc

period and I didn't do science at all for their third grading periods. Same for the social studies."

Teacher D: "…We don't take as many field trips. We don't do community outreach like we used to like visiting the nursing home or cleaning up the park because we had adopted a park and that was our job, to keep it clean. Well, we don't have time for that any more."

Teacher E: "I had to cut out some things in order to do the CSAP stuff. It's not the number of days. I think it would be more accurate to say the number of labs. I think what is more significant is that I have had to cut the number of hands-on investigations. I would say I have had to cut one quarter of the labs."

Teacher F: "Projects, [I] eliminated curriculum such as novels I would teach, we didn't have time to go to the library, we didn't have time to use the computer labs because they had to cut something. [I] Cut things I thought we could live with out. [I] Cut presentations, anything that takes very much time, I cut film. We have been cutting like crazy."

The following Florida teachers know what test pressure really feels like.[81] They seem to have had their professional judgment about curriculum dismissed.

Teacher A: "The FCAT is teaching teachers to stay within the narrow confines of the FCAT. Too many times I've been told, when going beyond the confines (especially in math): 'Why are you teaching that? It isn't on the FCAT.' "

Teacher B: "Our total curriculum is focused on reading, writing, and math. There is no extra time for students to study the arts, have physical education, science, or social studies. Our curriculum is very unbalanced."

This document is available on the Education Policy Studies Laboratory website at:
http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc

Teacher C: “While it is a way of testing some components of standards based performance, it leaves many gaps in the educational process. If we just ‘teach to the test’ which many teachers in our district are pressured to do, then the students are left with HUGE educational gaps that have not been covered in their education. Students deserve a well-rounded education, not just bits and pieces that are presented on a state test.”

Teacher D: “Before FCAT I was a better teacher. I was exposing my children to a wide range of science and social studies experiences. I taught using themes that really immersed the children into learning about a topic using their reading, writing, math, and technology skills. Now I’m basically afraid to NOT teach to the test. I know that the way I was teaching was building a better foundation for my kids as well as a love of learning. Now each year I can’t wait until March is over so I can spend the last two and a half months of school teaching the way I want to teach, the way I know students will be excited about.”

This document is available on the Education Policy Studies Laboratory website at:
http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc

Table 6: Narrowing the Curriculum

Location of Story	Source	Headline	Story
1. Florida	St. Petersburg Times, Rebecca Catalanello (Staff Writer) (February 1, 2004).	Kicking the 'FCAT essay' habit	Writing teachers lament the pressure of writing exams that dilute student creativity. The story notes, "From the time they start taking the FCAT writing test in the fourth grade, Florida students are taught how to navigate the 45-minute exam. That's good for scores, critics say, but often terrible for the emerging writer." One student talks about how the test forces her to subdue her creativity. "When she writes for pleasure, Stark strives for effect, atmosphere and subtlety. But when she writes for the test, she takes no chances. On Feb. 10, the next time the test will be given, she will do the best she can in five paragraphs, making three points and using the kinds of transitional words she has been told readers enjoy – "first," "next," "in conclusion." "It's like mechanics," Stark says. "I do what they want, I spit it out and then I move on."
2. California	Pasadena Star-News, Marie Leech (Staff writer) (October 13, 2002).	Teachers question testing frenzy	Teachers discuss the problems associated with an educational reform movement driven by standardized testing. One fifth-grade teacher says that "the state's rigorous testing schedule is failing students because it leaves teachers no choice but to base instruction on what's being asked on the tests." "Tests have become the holy grail of education," said Wiener, a teacher at Ynez School in Alhambra. Because he has to focus so much on math, reading and language arts, Wiener has had to "downplay many other subjects, especially the arts. "I think public speaking is important, but it takes up a lot of time and it's not on the test," he said. Weeks can go by without students having art and music classes, he said. Mary White, a third-grade teacher at La Seda Elementary in La Puente, said she doesn't have enough time in the day to have separate lessons for science and social science. Instead, she incorporates those texts into her reading lessons so students learn to read in different ways, White said.
3. Florida	The Stuart News/Port St. Lucie News, Editorial written by 17-year-old David Buckey, (July 1, 2002), p. A8.	Get serious about making education higher priority	As a 17-year-old senior-to-be at Martin County High School, it is clear that Martin County and the state of Florida do not give education the priority it should have. The politicians and residents of Martin County are so obsessed over whether to move the airport that they ignore some classrooms approaching 30 students, meager teacher salaries and teachers buying classroom supplies out-of-pocket. Instead, we are spending $150,000 on a noise study so that we can erect useless noise barriers that will cost more money. In addition, Florida instituted the FCAT test to see if students are learning what they are supposed to learn. The FCAT results in a narrow curriculum that forces teachers to cut out subjects in order to spend time on FCAT drill. The result of this is students unprepared for college or careers after high school.

This document is available on the Education Policy Studies Laboratory website at: http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc

Table 6, continued

4. Georgia	The Atlanta Journal-Constitution Editorial, (May 7, 2002), p. A21.	Students more than labels	An editorial decrying the policies of judging students and schools on test scores. The writer argues that before NCLB, Georgia had already tried to use standardized testing to diagnose weaknesses (not as a way to dole out sanctions and rewards as NCLB does). Back then, "Researchers pointed to strong evidence that overzealous testing narrows curriculum, stifles innovation and reduces instruction to rote drill." The writer concludes that we haven't made any progress and that NCLB and its use of standardized test scores to make decisions about schools only "leaves kids right back where they started – at their original schools, which now have been publicly disparaged as failing."
5. Memphis, Tennessee	The Commercial Appeal, Editorial Viewpoint (September 24, 2003), p. B4.	High-stakes testing turns the screw	Viewpoint citing the pressures on teachers and administrators to raise standardized test scores. This pressure has forced many teachers to alter their curriculum. An unscientific survey of teachers (asking how high-stakes testing is affecting them), sponsored by this newspaper, found that many teachers are having to narrow their curriculum. One middle-school teacher complained of being unable to teach children about America's first moon landing because of the demands a heavy testing schedule places on classroom time. A school's fortunes rise or fall according to the results of a multiple-choice test battery given once a year - Tennessee Comprehensive Assessment Program (TCAP) achievement tests for elementary and middle school students, Gateway tests for high school students. Defenders of high-stakes testing insist that drilling students on the kind of material that will be included on tests is not, strictly speaking, "teaching to the test," which would shortchange pupils with an excessively narrow curriculum, as critics charge.
6. National perspective	New York Times, Walt Gardner (July 6, 2003), p. 5.	Personal finance in the classroom	Writer taught for 28 years in the Los Angeles School District and complains about judging schools almost exclusively by standardized test scores. This pressure "narrows the curriculum" and "shortchanges students who would otherwise profit from more creativity."
7. National perspective	The Washington Times, George Archibald (July 3, 2003), p. A2.	NEA rejects outline by Bush for schools; says plan would have adverse effect	More than 1,000 teachers loudly applauded David Lebow, chairman of NEA's professional standards and practices committee, when he said at a forum, "There'll be no behind left" under the administration's federal No Child Left Behind Act, which set up a 10-year program of testing and reform measures to improve the reading and mathematics proficiency of all students. "Teachers are being forced to narrow the curriculum, pay less attention to critical thinking skills" because of new reading and mathematics testing requirements from third through eighth grade, said Becky Pringle, a member of the NEA's executive committee, who moderated the forum. According to the article, "The NEA, the country's leading teachers union, has sent Congress a list of 47 changes it wants in the law in order to give states and school districts more flexibility in meeting the measure's requirements."

This document is available on the Education Policy Studies Laboratory website at: **http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc**

Table 6, continued

8.Colorado	Greeley Tribune Lisa Martinez, (February 15, 2004).	CSAP rebellion brewing	A story about 22-year-old elementary education major Elena Mendoza who is on a mission to stop a standardized test that thousands of Colorado children take each year. She is part of a new coalition for REASON in education (REASON stands for Resistance Against Standardized Onslaught of Nonsense). Her cause is to persuade parents across the state to keep their child home when CSAP is being administered arguing that the pressure associated with the test is narrowing the curriculum. According to her, "CSAP takes away from learning in the classroom and focuses only on the tests." Mendoza talked with parents and teachers and made a presentation to the school board of one district. "One teacher she talked to said she went two weeks without teaching art or social studies because those subjects were not tested on CSAP… those students were only working on reading, writing and math." According to Mendoza, "I was appalled … we can't teach kids to learn. We teach them to take tests."
9. Maryland	Annapolis Capitol Gazette, Kimberly Marselas (June 26, 2003).	Fine arts panel warns of eroding arts program	Anne Arundel County Maryland lost 23 middle school art teachers.*
10. Oregon	Wall Street Journal, Anne Marie Chaker (October 30, 2003).	Schools say 'Adieu' to foreign languages	The Rosenburg Public schools in Oregon cancelled seventh- and eighth-grade foreign language classes.*
11. Arizona	ABCNews.com, Geraldine Sealey (August 25, 2003).	Just the three R's?	The Arizona legislature cut $7 million in arts funding to schools and other groups.*
12. Wisconsin	School Board News, Del Stover (December 16, 2003).	Budget cuts threaten core programs	Milwaukee, Wisconsin has lost nearly 9 percent of its art, music and physical education teachers.*

This document is available on the Education Policy Studies Laboratory website at: **http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc**

Table 6, continued

13. Rhode Island	National Education Association (December 2003/January 2004).	Cuts leave more and more public school children behind	Providence, Rhode Island eliminated elementary science and technology-enrichment classes.*
			*The above five references are taken from Claus von Zastrow (with Helen Janc) (2004, March) Academic Atrophy: The condition of the liberal arts in America's public schools. Council for Basic Education

This document is available on the Education Policy Studies Laboratory website at: **http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc**

This small table of news articles barely touches on the problem of a narrowed curriculum as a function of high-stakes testing. Not only are courses and areas being jettisoned, but within courses that are related to the tests a narrow vision of the curriculum is being fostered. The overall result of high-stakes testing, for many teachers and school districts, seems to be a restricted vision of what should be taught and how it should be taught. Less obvious but of equal concern is that there also exists a restricted vision of the kind of person who should be teaching. Implied is that a good deal of teaching in high-stakes testing environments is of a technical rather than a professional nature. Technicians and professionals are prepared differently, are expected to work differently, and have different obligations to those they serve. Here, in the voice of Ann, a new teacher, we see how a professional is destroyed and a technician is born:[82]

> Last year, when I was college student, I had great ideas using hands-on activities and cooperative learning in my classroom as a way to get students to be internally motivated for learning. With the testing programs we have in this school, there isn't much leeway for me to be creative or innovative and create excellent lessons. The test is the total goal. We spend time every day doing rote exercises. Forget ever doing hands-on activities, science or math games, or creative writing experiences. We do one hour of sit and drill in each of the subjects of math, reading, and writing. We use a basal reader, math workbook pages, and rote writing prompts. It is all step by step; the same thing every week. Every day for one hour the whole school does the exact same instruction lesson. No visitors are allowed in the classes. The children sit and get drilled over and over. I have to teach the letters by saying "A, what is A?" I repeat this over and over in a

This document is available on the Education Policy Studies Laboratory website at:
http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc

scripted lesson that my principal makes me use. I read exactly what they hear. You can't improvise, add, or take away. You read exactly what it says. This is how testing has impacted my school and my teaching. As a first year teacher I feel like I don't have a choice to deviate from this awful test preparation.

Conflicting Accountability Ratings

Table 7 presents articles demonstrating that different evaluation systems do not always give the same information. This is to be expected. But if evaluation systems are each examining the same construct, they should at least be correlated positively. The stories in this table suggest that oftentimes no such correlation exists. Article 10 makes this point as well as any others in the table. In that article we see a school dubbed "outstanding" by no less a person than our President, and then the federal NCLB system of school evaluation kicks in and the school is designated a failure. The discrepancy between local evaluations and those that that are required by NCLB is nothing short of confusing, as Albuquerque, NM, has discovered: [83]

> Consider the absurdities in the current New Mexico school rankings for Albuquerque:
>
> Desert Ridge Middle School scored 88 percent proficient in math and 89 percent in language arts in eighth grade. But it failed to achieve "adequate yearly progress."
>
> Madison Middle School scored 77 percent of students proficient in language arts and 69 percent in math. But it failed "adequate yearly progress."

This document is available on the Education Policy Studies Laboratory website at:
http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc

Jackson Middle School scored 76 percent proficient in language arts and 65 percent in math. But it failed "adequate yearly progress."

Griegos Elementary School scored 78 percent proficient in language arts and 92 percent in math. But it failed "adequate yearly progress."

Osuna Elementary School scored 83 percent proficiency in math and 74 percent in language arts. But it failed "adequate yearly progress."

On the other hand: Navajo Elementary School scored 29 percent proficiency in language arts and 32 percent in math. It achieved "adequate yearly progress."

Armijo Elementary School had 41 percent proficiency in language arts and 46 percent in math. It achieved "adequate yearly progress."

Eldorado and Cibola are among the best high schools in the state. They failed to achieve "adequate yearly progress."

How can such apparent discrepancies occur so frequently? The answer emerges from studying the stories in Table 7.

This document is available on the Education Policy Studies Laboratory website at:
http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc

Table 7: Conflicting Accountability Ratings

Location of Story	Source	Headline	Story
1. National perspective	New York Times, Sam Dillon (September 5, 2004).	Good schools or bad? Ratings baffle parents	As the fall, 2004 school year gets underway, reports emerged regarding the numerous schools across the nation that had received two ratings. According to the article, "In North Carolina, more than 32 schools ranked as excellent by the state failed to make federally mandated progress. In California, 317 schools showed tremendous growth on the state's performance index, but were labeled as low-performing according to the federal district."
2. Richmond, Virginia	New York Times, Michael Winerip (November 19, 2003).	Superior school fails a crucial federal test	The author describes the accolades of one Virginia middle school as the "kind of school any child would be lucky to attend. The district has given every middle and high school student a laptop computer, 27,000 in all. Tuckahoe's (middle school) test scores are among the best in Virginia, with 99 percent achieving proficiency in math, 95 percent in English. Its previous principal was the 2002 state principal of the year, and in 1996, Tuckahoe was named a Blue Ribbon School of Excellence by the federal Education Department." But in September 2003, Tuckahoe was labeled a failure under the federal system. In Tuckahoe, 94 percent (instead of the mandated 95 percent) of students were tested.
3. St. Petersburg, Florida	St. Petersburg Times, Barbara Behrendt (August 9, 2003), p. 1.	Local schools left behind by federal grade system	Just after school ended in May 2003, the state announced that every elementary and middle school in one district earned an A and every high school had earned a B as part of the state's grading plan. But, as of August 2003, most of those schools did not meet the federal requirements for academic progress and were labeled as "failing."
4. Florida	Miami Herald Mathew I. Pinzur (October 21, 2003).	Paradox in school bonus for FCAT Success	In Florida, roughly half the students at one school failed to meet state standards in math and reading during the 2002-2003 school year, yet the school received an A grade from the state and $175,208 in bonus funding. In another elementary school, 72 percent of students met standards for reading (63 percent for math) on the FCAT, but it got a C on the state system and received no funding.
5. Boston, Massachusetts	Megan Tench (Staff Writer) Boston Globe.	Conflicting evaluations puzzle school	Discusses the confusion of one elementary school, the largest in Boston, which is located in a crime-heavy and poverty-ridden intersection of town. The school has been publicly lauded for its success in achieving good reading scores. In fact, the states superintendent of schools announced the school as "one of the most progressive in the city two years in a row, a school that exhibits 'effective practices.'" Yet, over those same two years, the federal government guidelines decried the school as failing to make AYP.

This document is available on the Education Policy Studies Laboratory website at: http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc

Table 7, continued

6. Arizona	Arizona Republic, Mel Melendez, (October 22, 2003).	School ratings confusing parents	Article about how confusing dual ratings are to parents across the state. For example, one school after earning the highest "excelling" state ranking, had also failed to make Adequate Year Progress and was dubbed "failing" by the federal ranking system.
7. St. Louis, Missouri	St. Louis Post-Dispatch, Carolyn Bower (Staff Writer) (September 3, 2003).	Many area schools fail test: Some that fell short got recognition from state for performance	About half of Missouri's 2,055 schools, including most middle and high schools in the St. Louis region, did not meet standards in communication arts and math this year. Yet, some of these schools were among the top 10 performers on Missouri tests last year. Indeed, many districts that did not meet standards are the same ones the state recognized as outstanding, for performance on state tests.
8. Portland, Oregon	KATU News, Julia Silverman, (January 29, 2004); & Associated Press, Oregon Daily News (January 29, 2004).	Federal and state standards very different for schools. & Oregon report cards counteract federal grades for some schools	About 300 Oregon schools received a label of "failing" from the federal government. However, as of this week, many of these same schools received an "A" from the state. The state report card is a snapshot of the school's overall performance, while the federal rating is significantly affected by individual student group performance. One school made improvements in math and received an "exceptional" rating from the state this year. Yet, the school got a failing grade from the federal government because a handful of its developmentally delayed students didn't achieve grade-level proficiency.
9. Detroit, Michigan	The Detroit News, Christine MacDonald & Maureen Feighan (Staff writers), (February 3, 2004).	Making the Grade: Schools give selves A's, avert failure	Writers discuss how many of Michigan's chronically failing schools are giving themselves the highest possible marks on state report cards – saving themselves from a D or flunking grade that would have brought their students more oversight and assistance. For example, "Schools in Utica, Pontiac, Detroit, and Wayne-Westland were among the 76 percent of troubled schools statewide that gave themselves A's on a self-evaluation that was worth a third of their grades in report cards issues last week." The self-evaluation measures 11 factors, such as how well the schools "reach out to parents, their building's condition and how well they prepare their teachers." Notably, the self-evaluation does not ask about test scores on the MEAP. Experts are critical of the self-evaluation process that is too subjective.
10. Michigan	Detroit Free Press, Lori Higgins (Staff Writer), (February 11, 2004).	Top metro schools are labeled failures: Award-winners fall short of federal standards	Story about several schools across the state that have won major accolades for their high achievement. For example, "A Southfield school was hailed as a model by President George W. Bush. Other schools across Michigan received National Blue Ribbon awards. And others in some of the state's wealthiest districts are considered first-class high schools." Yet, one major commonality across these schools is that they are considered "failing" under NCLB. In fact, "some high-achieving schools even received 'F' grades on their school report cards, released late last month by the Michigan Department of Education, because their high scores aren't improving fast enough."

This document is available on the Education Policy Studies Laboratory website at: **http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc**

Table 7, continued

11. Eugene, Oregon	The Register Guard, Staff, (February 4, 2004).	Blunt instruments: School report cards can be meaningless	All over the state, schools rated by the state as "strong" or "exceptional" were labeled as failing to meet federal standards. The year-to-year inconsistencies and the lack of agreement between different grading systems, calls into question the entire concept of attaching a single grade to an entire school. But when nearly all schools are rated as failing, as occurred under the federal standards, the measure becomes meaningless. The federal standards require that schools meet a list of performance and attendance standards for all types of students -- the disabled, the economically disadvantaged, ethnic or racial minorities, non-English speakers. If any group of students falls short, the entire school is judged to have failed. Over time, repeated failure can cause a school to lose federal funds. A closer look would reveal exceptionally good things happening, as well as areas needing improvement, in every school. A program that yields strong results for one child might prove satisfactory or worse for another. The report cards, both federal and state, are blunt instruments for measuring school quality and obscure a true understanding of what is actually happening in classrooms. Parents can come up with more useful grades on their own, by talking to their children and to their children's teachers.
12. Billings, Montana	Billings Gazette, John Fitzgerald (Staff writer), (March 3, 2004).	Schools earn high rankings	For a second year in a row, West and Skyview high schools have earned the highest mark given by regional accreditation groups, but neither school meets federal No Child Left Behind standards. According to Skyview Principal Bob Whalen, "It's hard for us to figure." According to the West principal, Dennis Sulser, "By one organization we're succeeding, by another we're failing." In Montana, the Northwest Association of Schools Colleges and Universities has given West and Skyview its highest ranking – exemplary status. According to the article "The group gives accreditation to schools that meet its standards, but issue an exemplary designation to very few. Of the 13 AA high schools in Montana the association gave the status to only 5."
13. Missouri	Associated Press, Kelly Wiese (January 16, 2004).	Dozens of school districts both outstanding and failing	Chillicothe School District was recognized in the state for "distinction in performance" at the same time, it was found to be "failing" according to the federal No Child Left Behind Act. According to the report, this mixed message is not unusual. The report goes on, "In all, 153 school districts were recognized by Gov. Bob Holden last month for 'distinction in performance' for the 2002-03 school year. Yet 90 of those districts had schools that failed to make 'adequate yearly progress' on standardized tests in reading or math under the federal education law." Understandably, administrators find it tough to explain to parents how their children's schools are outstanding and failing at the same time. "It's hard to make that distinction," said Chillicothe Superintendent Dale Wallace. "It's difficult for people to understand."

This document is available on the Education Policy Studies Laboratory website at: **http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc**

Table 7, continued

14. Mississippi	Associated Press (September 5, 2003).	State schools fare better under state standards than federal & Some Mississippi schools excel in state system, fall short in federal	Under the state standards, nearly 83 percent of Mississippi schools were rated as successful or higher. However, only half of the state's schools met the new federal requirements. Of the 822 public schools that received state ratings, 150 were rated as superior-performing, and only 33 were considered low performing. South Forrest Attendance Center Principal Dale Coleman got some mixed signals this week when his school's rating was released under new state and federal accountability standards. The K-8 school in Hattiesburg is among Mississippi's top public schools, based on the state accountability system. But South Forrest failed to meet the federal No Child Left Behind standards. “I have one of the best schools around,” Coleman said Friday, of his school's Level 5 - or superior - rating by the state. “You don't get to be Level 5 not having your stuff in line.”
15. West Virginia	Charleston Gazette, Eric Eyre (September 25, 2003).	A twist to No Child Left Behind: West Virginia uses new standards to praise schools	State school administrators don't want the federal NCLB guidelines to tarnish every school in the state. So, they are doing a “Recognition Tour – going around and giving certificates to schools that made AYP, but also to schools that may have not made AYP but at least made some progress.” In July, the state released a list of 326 schools that failed to meet the federal standards. The state has 728 schools, so about 45 percent of schools didn't meet the mark. Ironically, some schools on the low-performing list had received state accolades as “exemplary schools” in previous years. Forty-two schools fell below standards two years in a row, meaning they might have to provide tutoring services or transportation for students who want to switch schools. The news about struggling schools traveled swiftly. Parents, teachers, principals, school board members and legislators were angered. The federal government had given their schools a black eye, but not the money to improve them, they said.

This document is available on the Education Policy Studies Laboratory website at: **http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc**

First of all, some of the schools were self-evaluating, meaning that the chance for bias to enter into the evaluation was quite high. Self-evaluations are too often self-serving evaluations and thus they would be expected to be different than those that are based on some outside criteria, like an achievement test. Article 9 describes such a situation.

School evaluations can produce different outcomes when districts or schools win awards based on "soft" criteria (award winning community service programs), earn good publicity (a student team winning an academic decathlon), host special programs (an award-winning choir), and so forth. Such schools often are admired in their community, receive commendations from their governor, and accumulate awards as A+ schools. But under NCLB such schools might not be making enough progress annually, and so they are labeled failures by people who do not know these schools and do not choose to recognize an alternative set of evaluative criteria.

The NCLB accountability system demands annual growth by many different subgroups of students, in many different grades, and thus it has built into it many ways to make a good school look like a failure, as is clear from the Albuquerque example. Table 7 also informs us that the different systems for school evaluation are not correlated highly and this makes it difficult for the public to know which is more valid—a community's judgment that their local school is superb, or the federal judgment that it is not. This dilemma is most evident in the story reported in Article 2. We believe that this difference in evaluation systems is an important issue for educators and bureaucrats to discuss over the next few years. From our point of view we conclude that there is merit to the local evaluations because we see schools as having many goals along with academic

This document is available on the Education Policy Studies Laboratory website at:
http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc

achievement. Thus the current federal NCLB evaluation system is seen as overly narrow and seriously, perhaps irredeemably flawed. That system needs to be modified to include a wider set of indictors for evaluating schools. Mere growth in the numbers of students who reach the proficient category, with each state having its own definitions of proficient, should not be the only yardstick by which school quality is measured.

But the states' own internal systems of evaluation are seriously flawed too, as seen in Articles 4 and 7. The latter article also illustrates the folly of using a single year's worth of data to describe a school as failing or succeeding, and rewarding or punishing it on that basis. Year-to-year score perturbations are expected to be substantial in most schools, but they will be especially large in small schools.[84] Thus *all* accountability systems for small schools that use annual testing for judging proficiency or growth are so seriously flawed as to be useless. School score gains from year to year are actually negatively correlated, so using such a system is sure to give us a different perception of a school from one year to another.

Mere growth in the numbers of students who reach the proficient category, with each state having its own definitions of proficient, should not be the only yardstick by which school quality is measured.

How does this table illustrate Campbell's law? It does so indirectly by showing that accountability systems yield different results and under such circumstances we should expect that educators will pick the system that makes them look the best. That should not be surprising. Thus, through these articles, we see a defense of local evaluation systems and attacks upon the federal system, which seems to rely upon a harsher indicator. But it is not clear to us that either evaluation system is noticeably

This document is available on the Education Policy Studies Laboratory website at:
http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc

better than the other. Good evaluations of schools are conducted over considerable lengths of time, and they rely upon multiple indicators of success. Neither federal nor local evaluation systems usually meet these criteria. And neither system of evaluating achievement would stand up to close scrutiny whether they are value-added or static models of student learning because our psychometrics are not up to the requirements that politicians place on our research community. When there exists no consensus about which indicators might be most appropriate to use in the evaluation of schools, Campbell's law suggests that there might be more support for the indicators that make schools and districts look the best, rather than engagement in a search to get a better indicator system. That should be guarded against.

The Changing Meaning of Proficiency

Table 8 presents stories of how the meaning of proficiency in various communities changes over time. In high-stakes testing, to pass and fail students, there has to be a point on a scale where one decides that students above or below that point have learned or not learned what is expected. Students either are or are not proficient; they either have earned the right to be passed to the next grade, or not; they either are worthy of graduation, or not. While this sounds like a reasonable problem, and that a reasonable solution might follow, it turns out that this is an enormously complex problem, and it has not yet been solved by some of the brightest people in the field of testing.[85] Gene Glass wrote on standard setting for tests almost 30 years ago. Standard setting means that you have determined some absolute standard of performance, above which someone is competent and below which they are not. An example might be that a

This document is available on the Education Policy Studies Laboratory website at:
http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc

score of 130 or more on a particular IQ test allows you into a program for gifted children, but a score of 129 or less keeps you out. Glass says:

> I am confident that the only sensible interpretations of data from assessment programs will be based solely on whether the rate of performance goes up or down. Interpretations and decisions based on absolute levels of performance….will be largely meaningless, since these absolute levels vary unaccountably with exercise content and difficulty, since judges will disagree wildly on the question of what consequences ought to ensue from the same absolute level of performance, and since there is no way to relate absolute levels of performance on exercises to success on the job, at higher levels of schooling, or in life. Setting performance standards on tests….by known methods is a waste of time or worse.[86]

Choosing cut scores on tests, that is, determining the absolute level at which bad performance turns miraculously into good performance, is simply impossible to do. Worse than wasting time trying to do it, is fooling people into thinking it can done. Stated as clearly as possible: The choice of the cut point for high-stakes achievement tests is arbitrary. It is a political decision, not a scientific one. By political we mean that cut scores are determined by the acceptable failure rates in particular communities. Typically, this means that wealthy white students, whose parents have political capital, pass at high rates, while poor minority students, whose parents have no such capital, fail at high rates. Cut scores for high-stakes tests in most states appears to be about choosing an acceptable level of casualties!

This document is available on the Education Policy Studies Laboratory website at:
http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc

Table 8: Changing Meaning of Proficiency

Location of Story	Source	Headline	Story
1. Portland, Maine	Portland Press Herald, Tess Nacelewicz (Staff Writer) (December 24, 2002), p. 1A.	Maine may ease education goals	After raising the achievement benchmarks, Maine considers lowering them in order to address a concern that the high benchmarks hurt the state when it comes to NCLB. The MEA was revised a few years prior to make it tougher. But now, the new scoring standard is higher than those for other states and puts Maine at a disadvantage when it comes to ranking schools.
2. Lexington, Kentucky	Lexington Herald-Leader, Lisa Deffendall (October 19, 2003).	No statistics are being left behind: State has wide leeway in grading schools	A new statistical adjustment for measurement errors on Kentucky's statewide achievement test (incorporating confidence intervals in determining AYP) dramatically improves failure rate from 72.1 percent last year to 38.8 percent this year.
3. North Carolina and South Carolina	New York Times, Ford Fessenden (December 31, 2003).	How to measure student proficiency	Two middle schools, one in North Carolina and one in South Carolina, received very disparate rankings and achievement results because the states varied in how they defined proficiency. States have set widely different standards for measuring students' progress under NCLB. For example, three quarters of children across the country would fail South Carolina's tough fifth-grade test, while seven out of eight would ace the third grade tests in Colorado and Texas.
4. New York	Buffalo News, Peter Simon (July 18, 2001).	Teachers oppose test-grade adjusting	The state had decided to adjust student grades on the high-stakes Regents exams. Several Regents exams contain generous curves; the state describes them as predetermined scaled scoring adjustments. To get a passing score of 55 on last month's biology/living environment exam, students needed to earn 28 of 85 possible points – or just 33 percent of the material tested. To receive a score of 65 – the passing grade in many districts – students needed to correctly answer 46 percent of the material.. "Why would you design a test and say, 'You really only need to know 35 or 40 percent of it?'" Philip Rumore, president of the Buffalo Teachers Federation, said Tuesday, "It's like something from a bad Kafka novel." Several local teachers charged that the state is trying to avoid widespread student failure while trying to maintain the appearance of tough new standards.
5. New York	New York Times, Karen Arenson (August 30, 2003), p. B3.	Scores on Math Regents exam to be raised for thousands	The extremely low passing rates on the math exam prompted the state department to re-score it, allowing thousands of students who previously failed it to pass. The re-scoring means that 80 percent of ninth graders (versus 61 percent from before) will pass.

This document is available on the Education Policy Studies Laboratory website at: http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc

Table 8, continued

6. New York	New York Times, Karen W. Arenson (October 9, 2003).	New York to lower the bar for high school graduation	New York State's education commissioner announces that the state would loosen the demanding testing requirements in place for high school graduation, including the standards to judge math proficiency. In June the results on the math Regents exam for 11th and 12th graders were put aside (only 37 percent passed, whereas 61 percent passed the previous year).
7. New York	New York Times, Elissa Gootman, (October 26, 2003).	How fourth and eighth graders fared on the New York State Math Test	Following last year's results on the statewide Regents exam in math, where large percentage of students failed, this year's results suggest that the test may now be too easy. In the state, the proportion of eighth graders who met math standards grew by 10.5 percentage points, but in the city, it grew 14.7 points.
8. New York	New York Times Elissa Gootman, (Staff Writer) (January 10, 2004).	Thousands pass regents test under revised scoring	Article about how thousands of students who had thought they failed the Regents Physics exam had now passed because the exam was re-scored. Superintendents and principals had called for the re-scoring after test scores plummeted in 2002. Thirty-nine percent of students failed the physics test in 2002 leading educators to believe it was too hard … this was a higher than normal failure rate. In October of 2003, a state Board of Regents panel determined the test was too hard after a second year of high failure rates (in 2003, 47.1 percent of students failed it).
9. New York	The Timesunion.com Gordon E. Van Hooft, (February 15, 2004).	Regents exams are questionable	According to the report, one Buffalo teacher stated that "after the Math A exam given in January turned out to be too easy and raw scores were scaled so those who got only 28 percent correct of a possible 84 questions passed at 55, that 'from a moral and ethical standpoint, we're giving kids credit for things they should not be getting credit for, and the kids realize that.'" The problematic nature of the Regents Math exam is that either too many, or too few students pass – leading to controversial re-scoring decisions. According to the report, "in recent years, the results on other Regents exams, such as the biology, physics, and the Math A exam, have reflected the errors associated with setting the passing grades either too high or too low." Currently, New York is delaying a decision whether to raise the passing score of 55 to 65 on the Math A exam because it is still unclear how to handle the large numbers of students with marks between 55 and 65.

This document is available on the Education Policy Studies Laboratory website at: **http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc**

Table 8, continued

10. New York	The Independent Online (indenews.com), David Riley (May 4, 2004).	Regents exams face new test.	At the crux of the issue is whether or not state assemblymen should change the bar of proficiency for obtaining a high school diploma in New York State. Currently, students must pass five Regents examinations in five subject matter areas to graduate high school. The current debate over what the graduate requirements may look like stems from a series of hearing that the legislators held across the state. According to Republican Sate Senator Steve Saland, "Overwhelmingly, school officials asked for greater flexibility. Others were concerned that getting test results sometimes outweighs what should be the main goal in schools giving a good education." The article goes on, "So-called high-stakes testing may also be unfair to students. 'Some people simply learn differently … some students, for whatever reason, may not perform particularly well on a given day,' according to Mr. Saland." Last year the Regents had extended a "safety net," allowing students to receive local diplomas if they scored between 55 and 65 on the Regents examinations (a 65 currently is the passing mark needed to obtain a high school diploma). According to one teacher, the problem was not about flexibility but about letting educators play a bigger role in the testing process. "Teachers, for example, used to be widely surveyed to submit questions for upcoming state tests. While some teachers are still asked for input, much test-writing is done by private contractors today. Ms. Fox, a New York State teacher, argues, "The challenge is in coming up with a fair test."
11. New York and Chicago	The New York Post, Online Edition, Carl Campanile (March 25, 2004).	Windy City Schools let up.	In New York, third-grade students must pass a test of minimum proficiency in both math and reading to be promoted to the next grade. In contrast, Chicago Board of Education recently revised their promotion standards in light of thousands of students who were forced to repeat a grade since the high-stakes testing policy was implemented. In Chicago, they had imposed, arguably, the strictest promotion standards of any of the nation's major cities in 1998, requiring students to score well on national exams in both reading and math. However, on March 24, 2004, the Chicago Board of Education revised the policy by dropping the requirement that students perform well on the math test. Under the new policy, a student will be promoted if he does well on the reading test even if he or she flunks the math test. In Chicago, the promotion policy had applied to students in third, sixth and eighth grades.
12. Boston, Massachusetts	Boston Herald, Elizabeth W. Crowley & David Guarino (Staff Writers) (May 24, 2003), p. 1.	Free Pass: Thousands who flunk could get diplomas	Article discusses how the state house filed a motion to "weaken" the high-stakes MCAS test, pushing to grant thousands of students who have failed it a free pass to graduate. The mandate would grant a diploma to 4,800 seniors in bilingual, vocational, and special education programs who were denied a diploma in spring, 2003.
13. Tampa, Florida	Tampa Tribune, Marilyn Brown (September 19, 2003), p. 1.	Florida Miscalculates schools' federal marks	Six weeks after the Florida State Department of Education gave most schools the bad news; it reversed the labels for many. Sixty failing schools were now considered making AYP, and fifty nine schools previously believed to be making AYP were now labeled as failing.

Table 8, continued

14. Miami, Florida	Miami Herald, Mathew I. Pinzur (October 29, 2003).	Hundreds of third-graders promoted	Hundreds of third graders in Miami Dade County that initially failed the state's standardized reading exam are now eligible to be promoted after passing an alternate version of the exam. Still, this only represents 5 percent of the 6,622 students who were held back. In Florida, students who score in the top half of the national SAT-9 exam, but who fail FCAT are eligible to go onto the next grade. Some argue the high numbers of students who were held back were simply an outcome of a high cut off score.
15. Arizona	Arizona Daily Star, Sarah Garrecht Gassen & Jennifer Sterba (September 3, 2003).	State talks tweak for AIMS test	Troubles with the Arizona Instrument to Measure Standards (AIMS) have prompted state officials to change the AIMS test – on claims that it is simply too hard. In the most recent round of testing, most eighth graders across the state failed the math and writing portions (eight out of ten failed math, and over half failed writing portion).
16. Seattle, Washington	The Seattle Times Linda Shaw (April 14, 2002), p. B1.	Seventh-graders sing WASL blues	A third of fourth graders who passed the reading and math sections in 1998 failed those subjects as seventh graders three years later. This dismal showing alerted officials to question the adequacy of the test. Some argue that the bar is set too high. Others believe the WASL needs adjusting.
17. Seattle, Washington	Seattle Times, Linda Shaw (Staff Reporter) (January 27, 2004).	Prospect of dismal test scores renews debate over WASL	Scores of the latest round of statewide testing (Washington Assessment of Student Learning – WASL) were about to be released amid widespread concerns that thousands of high school students might not graduate because they failed the test. Questions are being debated about what to do with the test, and the State Board of Education is debating whether the exam is fair? Too hard? A valid measure of learning? And, whether it is reasonable to require next year's freshmen to pass it before they graduate in 2008.
18. Georgia	Atlanta Journal-Constitution, Mary Macdonald (September 17, 2003).	Failure rate on Gateway exam rises sharply	The percentage of high school students who failed the Gateway exam on their first attempt shot up dramatically this spring – with the worst showing in science. 22% failed the science section, a near tripling of the failure rate in a single year.
19. New Orleans, Louisiana	Times-Picayune, Mathew Brown (August 22, 2003), p. 1.	State raises bar on LEAP standard	They are predicting that twice as many Louisiana fourth graders (an estimated 14,000 students), will fail next spring's high stakes LEAP exam because officials raised the bar of the exam – believing it to be too easy. The change means fourth graders (and eighth graders in 2006) will have to score "basic" or above on either the math or English portion of the test and at least "approaching basic" on the other portion. Previously, it was required that students obtain at least an "approaching basic" on both sections.

This document is available on the Education Policy Studies Laboratory website at: **http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc**

Table 8, continued

20. Bangor, Maine	Bangor Daily News, Associated Press (March 4, 1995).	Teachers, administrators criticize latest MEA tests	A barrage of educators criticized the recent administration of MEAs as being too difficult for fourth graders. One question asked fourth graders to write about their mayor – a task made difficult for at least seven towns who don't even have mayors.
21. Bangor, Maine	Associated Press (June 30, 1995).	Complaints prompt state to change achievement tests for 4th graders	State officials decided to change next year's fourth grade achievement tests to make it easier for Maine's 9- and 10-year olds. The Department of Education opted to change the Maine Educational Assessments (MEA) because teachers said the questions were written well above the fourth grade level, creating anxiety among students. Some of the changes include the following: 1) The tests, which were held every day for a week the previous year, will be broken up into shorter sections over a two-week period. 2) Students will write their answers in the same book. The previous year they had to record their answers in two books, making it confusing for students. 3) Last year the tests had no multiple-choice sections. In prior years at least half of the questions were multiple choice. The test the next year will not have any multiple choice questions.

This document is available on the Education Policy Studies Laboratory website at: **http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc**

The overwhelming conclusion drawn from these stories of changing proficiency levels is that parents, business people, politicians and educators have no defensible way to define proficiency. There simply is no agreed upon way to determine a test's cut score.

Article 1 documents the competitive nature of these cut scores, as well. If state A chooses a score to define proficiency that is higher than state B, then state A will fail more students and the educational system will look to be less effective. As noted in this story, that problem can be fixed by having state A drop their cut score lower, thus making itself look better. Under these circumstances the indicator of student achievement is not at all trustworthy, another instance of Campbell's law at work.

The overwhelming conclusion drawn from these stories of changing proficiency levels is that parents, business people, politicians and educators have no defensible way to define proficiency.

Article 2 and others in this table make the point that cut scores generally (but not always) go down. This is so a politically acceptable solution to the cut score problem can be found and more students can pass. Less frequently the reverse is true, where too many students appear to be passing, so a tests' cut score is raised (see Article 9). The more typical danger in all this dancing around a cut score for a high-stakes test is that after the politics plays out, the tests might resemble minimum competency tests. Such tests might ultimately allow almost all students to pass, thus diminishing the assumed benefit of high-stakes tests, namely, to motivate students and teachers to work harder. Because of politics there is not only pressure to lower cut scores, there is also a tendency for the test content to decrease in difficulty over the years, or for the scores to be allowed to drift

This document is available on the Education Policy Studies Laboratory website at:
http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc

upward as teachers and students learn what is on the test and teach to it more directly. In all three cases passing rates go up and the indictor no longer means what it once did. It becomes corrupted.

Articles 4 to 11, all concerning the state of New York, show that New Yorkers cannot agree at all on whether their tests are too hard or too easy. Moreover, they do not know what to do about it. They do not know what to do because there is no way to get unanimity among those concerned with the choice of a cut score. Political agreements must be reached. The technical skills of the psychometricians who made the test are not useful in making these kinds of decisions, as was discovered when state designations of proficiency in reading were evaluated against federal designations of proficiency. In this report for the Carnegie Corporation of New York, by Rand Corporation of Santa Monica, incredible variations between the federal and states' visions of proficient were revealed.[87] For example, Texas claimed that 85 percent of its students were proficient in reading at the fourth grade, using its own tests. But the federal government claims that only 27 percent of Texas' students are proficient in reading using the National Assessment for Educational Progress (NAEP) test. Mississippi claimed 87 percent proficiency in reading, but the NAEP data suggested that only 18 percent should be declared proficient. Only a few states showed any close relationship between its reading proficiency designations and those derived from the federal government's assessments. But in not a single case among the 40 states in the study was there a lower rate of proficiency than claimed by the federal government. States always had more students proficient in reading than suggested by their NAEP scores. Which exams might be used for judging schools, districts and states? No one knows because no one knows where to place a cut

This document is available on the Education Policy Studies Laboratory website at:
http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc

score on any examination. Moreover, there is considerable evidence that the NAEP Governing Board has deliberately set its cut scores to make it very difficult to reach proficiency. The National Academy of Sciences panel looking into standard setting on NAEP said, "NAEP's current achievement level setting procedures remain fundamentally flawed. The judgment tasks are difficult and confusing; raters' judgments of different item types are internally inconsistent; appropriate validity evidence for the cut scores is lacking; and the process has produced unreasonable results."[88] But because the states' systems of setting standards are faulty too, no one knows how to judge proficiency on their tests in a way that is convincing!

Cut scores are only one problem in determining who is declared proficient. Another problem is associated with the content of the tests, as made clear in Article 20. Arizona has its own example of this problem. From interviews with teachers of young children in Arizona, Annapurna Ganesh learned that there is an item on the SAT9 test about a blizzard.[89] Flagstaff students who experience winter might have gotten this item right, but Phoenix students apparently did not. Ganesh also heard about another problematic item. This one showed a picture of a bus and asked where it was going. Two of the three choices were "school" and "grocery store," with the former choice being the correct one. But in this poor neighborhood, students rode buses to grocery stores and had no need of school buses because they lived close to the school. Needless to say, most students got the item wrong. What is troubling is that fixed cut scores, arbitrarily chosen, could result in some students being judged proficient and others being failed due to these ordinary and minor forms of bias that are inherent in all tests. It may not be possible to

This document is available on the Education Policy Studies Laboratory website at:
http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc

ever rid tests of all forms of bias, but it is possible to adopt more flexible criteria for deciding who is or is not proficient.

Harlow discovered another problem with fixed and arbitrarily chosen cut scores for high-stakes tests.[90] She interviewed eighth-grade students in New Zealand who had taken the science part of the TIMSS tests. The students were interviewed about a sample of items answered correctly and incorrectly. The interviews probed the students' understanding of the questions and the students' spoken explanations were scored. Comparing the written test results to the interview test results revealed large differences. Many students went up dramatically in the number of items they got correct. However some students received lower scores as a result of the assessment through interview than they did with the assessment by means of the paper and pencil test. The lesson to be learned from this study is that the choice of method and the choice of cut score interact. Some students would be declared proficient under one method of assessment and fail under another. So how can a cut score be anything but arbitrary? This work parallels that done by Shavelson and his colleagues who demonstrated that the inter-correlations of students' performance on three different assessment methods were only in the .30s.[91] Such low inter-correlations inform us that students judged to be proficient via one method of assessment could be judged a failure by another method, and vice versa. The choice of a cut score for a method of assessment is always an imperfect, highly politicized means for identifying those who pass and those who fail.

The choice of a cut score for a method of assessment is always an imperfect, highly politicized means for identifying those who pass and those who fail.

This document is available on the Education Policy Studies Laboratory website at:
http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc

The Morale of School Personnel

Table 9 presents numerous stories informing us that teachers and administrators are quite unhappy with their profession since high-stakes testing became such an integral part of schooling in America.

This document is available on the Education Policy Studies Laboratory website at:
http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc

Table 9: The Morale of School Personnel

Location of Story	Source	Headline	Story
1. Rock Hill, South Carolina	The Herald, Erica Pippins (May 4, 2003), p. 1A.	State exam carries high stakes for teachers and pupils alike	Article talks about how teachers and students alike have been stressed in anticipation of taking the statewide test the PACT (Palmetto Achievement Challenge Test) … believed to be one of the more rigorous test in the nation. According to one teacher, "It's just mind boggling to know that for two weeks, I am putting what I do in the hands of a 12-year old. I remind them that it's not just their name on the test. What they score is a reflection of me and the school." Another teacher repeated this concern: "That aspect unnerves teachers somewhat because we know every child is not going to be proficient, but they can be the best they can be. Still, I feel solely responsible for their test results."
2. Massachusetts	Daily Hampshire Gazette (June 28, 2000).	Tears and joy fill one school	Reports on teachers who are lamenting the growing pressures of the job. One teacher is taking early retirement in part because of the pressures associated with the state's MCAS test. "The MCAS is one of the reasons I'm leaving, because of the pressure it puts on the kids and everyone else. And, we don't even have the right materials to teach for it."
3. North Carolina	The Charlotte Observer, Debbie Cenziper, (August 12, 1998).	North Carolina teachers denounce ABCs testing, poll says	Story about statewide survey results that show that nearly half the teachers questioned would consider changing schools if theirs was designated "low performing" by the state. About half said they spent more than 40 percent of their time having students practice for end-of-grade tests. According to Gail Jones, associate professor of education, "We knew that a lot of teachers seemed to have a lot of concerns about the ABCs accountability program, but we were surprised at the depth of the degree to which they felt negatively." She goes on to say, "They have indicated that their job is much more stressful, much more."

This document is available on the Education Policy Studies Laboratory website at: http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc

Table 9, continued

4. New York City	Daily News, Kathleen Lucadamo (Staff Writer) (June 21, 2004).	Teachers 'fraid of third grade	Story about the "ballooning" numbers of third-grade teachers who are "begging for different assignments this fall, saying the tough new promotion rules for their students have dulled the joy of teaching." According to Michael Kerr who has taught at PS 192 for six years, "I would certainly never teach third grade again in New York City." The writer notes, "Kerr was one of a dozen teachers, principals and education experts who told the Daily News that the mayor's take-no-prisoners promotion policy has teachers fleeing from a once-plum assignment." For the first time this year, a single test determined whether third graders would be promoted to fourth grade. As a result teachers said they dedicated upwards of four months to drilling students during the school day and after school and on weekends. By mid-April, exam time, many teachers "were burned out and terrified that their test-weary students would fail." Nearly 12,000 did. "I can't take this stress," said Dawn Renta, a third-grade teacher at PS 149 in Jackson Heights, Queens. Renta secured a spot as a second-grade teacher next year. "Second grade is by far the most popular grade because the students are young and there is no testing. Third used to be popular but that isn't the case anymore," a Brooklyn principal said.
5. Denver, Colorado	Denver Post, Monte Whaley (Staff Writer) (August 24, 2003), p. B1.	Colorado schools battle attrition at top: Record 30 percent of superintendents in state new this year—job pressures cited	Article discusses a record number of school superintendents starting the '03-'04 school year in light of surging numbers of superintendents who left office. The pressures put on schools and districts to perform well on the annual CSAP exam was cited as one reason for the increasing attrition rate.
6. Houston, Texas	Houston Chronicle, Salatheia Bryant (March 12, 2001), p. A13.	TAAS-ing and turning: Students aren't the only ones losing sleep over tests	Story about how one elementary school principal suffers from sleepless nights and butterflies in her stomach every spring. The previous year, she was among 76 percent of Houston school districts that posted gains on the TAAS. As the spring 2001 set of tests approach, the principal will once again endure sleepless nights. This year, the principal is concerned about how the number of late transfers into her school from overcrowded HISD schools will affect its performance. As late as the previous week, she had eight new students.
7. Illinois	Chicago Tribune, Tracy Dell'Angela (Staff) (October 31, 2003).	School says exam leaves unfair mark: Controversial test brings 'failing' tag	Story about the diminished morale of teachers and administrators at Rollins School in Aurora who, despite working hard to help students achieve, are publicly humiliated by labels that their school is failing. The article notes that the label is particularly painful because the school has managed "to pull off a few test score miracles this year with one of the suburbs' most vulnerable student populations – a cash-strapped district where nearly every child is low-income and minority. The writer notes, "The failing label is crushing to staff morale and threatens to draw resources from the programs that have fueled Rollins' turnaround. And, the fact that it was caused by a test that never was intended to measure academic achievement makes the label all the more painful."

This document is available on the Education Policy Studies Laboratory website at: **http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc**

Table 9, continued

8. Portland, Maine	Portland Press Herald, Tom Bell (Staff Writer), (November 2, 2003), p. 1A.	Making the grade: Do 'report cards' on schools offer incentive, or harm?	Comments on the concerns teachers and principals have over the public ranking of schools, mostly using high-stakes but one-shot tests. One principal was quoted as saying of the practice of publicly ranking schools, "It's harmful. It creates panic. It's discouraging." Ironically, this principal led a school that was considered one of the state's best. However, this principal's statement is based on the fact that a school just across the river was ranked as failing. Article debates merits of public rankings, citing advantages of giving the public information to know how their schools are doing and disadvantages of how it harms teacher morale.
9. Maryland	Maryland Gazette, Editorial (Patricia Delawder) (March 13, 2002), p. A14).	MSPAP costs far outweigh the benefits	Editorial arguing that the cost of implementing the statewide MSPAP is not worth the benefits. "We are losing some of our best teachers and principals because they become frustrated with low scores and the intense pressure to move those scores up. Even our best teachers cannot make a child learn when he or she is not developmentally ready."
10. Rock Hill, South Carolina	The Herald, Erica Pippins (May 4, 2003), p. 1A.	Parents, communities, pitch in to help children study for PACT	In recognition of the increasing pressures being placed on teachers, community parents sponsored a workshop to help students prepare for the state test (PACT). According to one parent, "Teachers are stressed enough because they have to get so much instruction in during the day."
11. Portland, Oregon	KATU News, Julia Silverman (January 29, 2004).	Federal and state standards very different for schools	In Oregon, many schools that received a failing grade by the feds also received high grades from the state. The failure label from the feds affected teacher morale. Rocky Miner, principal at Crook County Middle School, said getting a "needing improvement" rating from the feds because of special education students and marginally lower attendance levels was tough on teacher morale this year. "But now, being rated 'strong' by the state, it helps people see that the hard work they have been doing is paying off," he said. "Our math scores over the past two years have improved 20 percent."
12. Annapolis, Maryland	The Capital, Leslie Gross, staff writer, (May 15, 1998).	Pep rallies kicked off the week in many county schools, while pizza and ice cream parties will celebrate its ending	It has nothing to do with sporting events, however. It is testing week. The special treats and other incentives are designed to lure children to school to take Maryland's statewide exam (MSPAP). For up to two hours every day some 17,000 third, fifth and eighth graders have been enduring the rigorous standardized exam. However, teachers are frustrated about the exam. According to one elementary teacher, "Two hours every day for five days? You're asking an awful lot." The article goes on to say that teachers complain about the extra work involved in preparing for the exam and the pressure on them to get their students to do well. One first-grade teacher said, "Between the pep rallies and daily cheers, there's too much pressure placed on teachers and students."

This document is available on the Education Policy Studies Laboratory website at: **http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc**

Table 9, continued

13. Michigan	Local 6 News, Jane Aldrich, (February 16, 2004).	MEAP Test controversy	In Michigan the pressure to perform on the statewide assessment has grown increasingly over time. Indeed, the Michigan Education Assessment Program has come under attack from parents, lawmakers, and even the teachers who give the test. Teachers are feeling the stress of the test. According to Kelly Kitchen, one eighth-grade teacher, "I think we all feel the stress of taking the MEAPs. Not only is it a test, but there's so much pressure on school districts to perform." One principal went on to say, "We're constantly having to prove to someone that we're good enough for their kids, but the measure that's being used is one or two tests. That causes stress for me, because we do so many things day in and day out that makes this a very good place for people to send their children." The article continues, "But what they do every day at Dwight Ridge Middle School doesn't matter when it comes to the MEAP. Principal Davis knows that." He goes on to say, "We accept that it's not fair. We still have the same task – to do the best job we can as educators with the kids." Another principal chimes in about the assessment system. She says, "Originally, it wasn't to compare schools." Principal Marge Clay has been a principal at Cornell Elementary school for 13 years. "She says in the past, tests just let you see how well students were doing." Marge Clay goes on to say, "There wasn't all this pressure to compare schools (before) No Child Left Behind, and now there's so much writing on how well your child does on MEAP."
14. Greensboro, North Carolina	News and Record, Editorial Letter to the Editor (October 13, 1995) p. A14.	Schools will fail if alternatives not provided	"As a recently retired teacher, I read with interest your article on the Professional Practices Commission's study of teacher turnover. Teachers do not flee their classrooms because they lack expertise, but because they feel unbearable frustration and stress. Under great pressure to raise test scores, teachers feel more accountable for student progress, while effective classroom management has become less and less achievable. Teachers must deal daily with increasing numbers of troubled, disruptive students who suffer the same ills which torment current society: poverty, abuse, values that glorify violence, drug and alcohol use, and criminal acts that go largely unpunished. These young people indeed need help, but no teacher, however well trained, can provide that help while simultaneously instructing 25 to 30 other would-be learners. Obviously, learning diminishes as disruption accelerates."
15. Greensboro, North Carolina	The News and Record, John Newsom, staff writer (May 20, 1997), p. A1.	Guilford County teachers are having test anxiety now that state test scores directly affect them	This story talks about the first administration of the three-day battery of tests that culminate a year's worth of schooling and weeks of review. Although students have taken these tests for years, the results carry greater weight this time around. For the first time, in 1997 the state will hold each school responsible for insuring its students learn everything they are supposed to do during the year. Teachers will get cash bonuses of up to $1,000 apiece. Schools where students fall short of state standards, state education officials may take control of the campus and teachers and principals could lose their jobs. According to the presidents of the Guilford County Association of Educators whose members include about two-thirds of the county's teachers, "Folks have been under stress, absolutely … The threat of people coming in and taking over your school, that's a little scary."

This document is available on the Education Policy Studies Laboratory website at: **http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc**

Table 9, continued

16. North Carolina	Charlotte Observer, Debbie Cenziper, (August 12, 1998).	North Carolina Teachers denounce ABC testing, pulses	In a survey sponsored by UNC Travelhill, North Carolina's elementary school teachers largely denounced the statewide testing program. "Morale has sunk, practice tests are soaking up teaching time, and students are more anxious about school than ever before." According to this survey, almost 70 percent of teachers believed the statewide testing program would not improve education at their schools, and many reported "that their colleagues have developed negative attitudes toward low achieving students." According to an associate professor of education, "We knew that a lot of teachers seemed to have a lot of concerns about the ABC's accountability program, but we were surprised at the depth of the degree to which they felt negatively." She goes on, "They have indicated that their job is much more stressful, much more. And in a time when they are projecting severe teacher losses in North Carolina and we already have problems recruiting and keeping good teachers, rather than running them off, I feel like we ought to be supporting them."
17. Bangor, Maine	The Bangor Daily News, Ruth-Ellen Cohen (Staff writer) (December 30, 2002), p. A1.	"Priority" Schools hit the books: Maine educators find opportunity in "failing" listing	The story is about several schools in the Bangor area. Among these were some schools labeled as "needs improvement." According to educators across the state, "The needs improvement designation has taken its toll…" "Our teachers work very hard and long hours – (the label) 'priority schools' damages staff morale." This is according to Jody Henderson, the principal of one school in Bangor. This principal contends that the label is bound to hiring repercussions aggravating the effects of the teacher shortage. She notes, "Who wants to work in a priority school when the hours are longer and the work load is heavier?" Paula Smith, principal at Pembroke Elementary School, "said the school's eighth-grade teacher told her she's never had a group take the MEAs so seriously as students did this year. They just poured their whole heart and soul into (the tests) because we explained to them: This is where we stand, this is where we are. It puts more pressure on them." The article concludes that in the end even schools with better labels are not immune to the pressures associated with the labeling system.
18. Portland, Maine	Portland Press Herald, Selena Ricks (Staff writer), (October 25, 2003) p. 1A.	Scarborough frustrated by low-performing label: the high school scores above the state average on MEAs, but not enough juniors took the exam	In Scarborough High School the scores on last year's MEA tests were higher than the state average in reading, writing, and math. But, according to the state, the high school is one of 142 public schools in Maine failing to meet NCLB guidelines for reading and math proficiency because not enough juniors completed the exam. The onus of the label was felt among the staff. The teachers were bothered to read news stories that their school is one of the "failing schools." According to one teacher, "To get upset, to internalize it would be counter-productive, but it bothers me what people are going to think about our school." A French teacher at the high school complains that, "It's not fair for the state to measure the performance of local schools by looking at only one test." The teacher goes on to say, "I know from young teachers to veterans like me that we are working hard to meet the needs of students and the students are working hard, too."
19. Florida	The Sun-Sentinel, Michael Mayo (March 23, 2004).	FCAT has a place, but not as be-all, end-all	Michael Mayo writes an article decrying the use of FCAT from making many high stakes decisions about students' promotion and retention. The pressure is affecting teachers and administrators as well as students. He writes about the test, "It has become the axis the school year is built around, spinning out of control. It has beaten the spirit from too many dedicated teachers, made administrators myopic and inflexible, and has soured kids on school when they should be at their most curious and enthusiastic."

This document is available on the Education Policy Studies Laboratory website at: **http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc**

Table 9, continued

20. California	California Educator, Posnick-Goodwin, Sherry) Volume 8, Issue 4 (December 2003).	Bring back the joy of teaching and learning	Story about the stresses of high stakes testing and their effects on students and teachers. For example, in one anecdotal story, Marie Avera came to the realization that the pressures of testing were forcing her to alter her method of instruction. The article notes, "Stressed out about raising test scores, the Vallejo City teacher was trying to drill a math concept into her students' brains when it became obvious that one child just didn't understand. "I said, somewhat sharply, 'Don't you get it?' He just looked at me with his big brown eyes. I felt terrible." At that point, she told her students to close their books. "'It's time for us to go out and play.' And I thought to myself, 'What are we doing to our children?'" CTA President Barbara E. Kerr says the biggest challenge facing schools is that the joy of teaching and learning is hard to find in today's classrooms. "Testing mania has gotten way out of hand and the Elementary and Secondary Education Act further exacerbates the problem," says Kerr, who taught first grade and kindergarten in Riverside for more than two decades. Everywhere she goes, teachers share the same concerns. In growing numbers of schools, scripted learning programs take up most of the school day, and the rest of the day is lost to pre-testing and testing. As for the joy: "It's as if someone is trying to throw a wet blanket over it every time you try and grab it. You see the glint in the eye of a kid after a great lesson, and you feel like you changed a life. No one can take that away from you. But we aren't given the chance to revel in that; we aren't encouraged to find more of those moments, because we are under pressures that have nothing to do with what is best for children."
21. California	Mike Perrault (Staff writer), The Desert Sun (March 22, 2004).	Judging test performance a tough task	The story documents the state's accountability system and how the school ranking system affects parents, students, teachers, and administrators across the state. Schools get an API score between 200 and 1,000, which can help show how they compare to one another. A different, 1-to-10 statewide ranking indicates how individual schools compare throughout California. A rank of 1 means the school scored in the bottom 10 percent schools statewide." The affects of this accountability system are being felt by teachers and administrators. For example, according to Michael Rosenfeld, a long time history teacher at a local high school, the ranking system "adds to the stress level – like a lot." He goes on to say, "I ignore test scores completely, 100 percent. I couldn't care less. All I can do is get out there and do my best." The teacher goes on to say, "Because there's so much emphasis, so much importance, placed on the test, you just have to give up a lot of other things that you think are important." The writer then says that the push for high rankings across the state of California is so high stakes that many administrators increasingly fear for their jobs. The history teacher goes on to say, "The way things are set up, it's their jobs that are on the line. They get a lot of pressure from the community, parents, the press."

This document is available on the Education Policy Studies Laboratory website at: **http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc**

Table 9, continued

22. California	The California Educator, Volume 5, Issue 5 (February 2001) Available online at http://www.cta.org/californiaeducator/z5i5/feature_1.htm, downloaded March 18, 2004	In 1999 California State Department adopted Senate bill 1X, the Public's School's Accountability Act of 1999	This act was intended to hold public schools accountable for the academic progress and achievement of pupils. This law established a fund to be granted to low performing schools. This fund was available under the Immediate Intervention/Underperforming Schools Program (II/USB). Under this program, schools that served students with low achievement and schools that have not shown academic progress for two years or more are eligible to apply for $50,000 for a startup fund to establish a planning progress. However, there is evidence that teachers in schools that apply and receive these funds experienced heightened sense of pressure and a decreased sense of morale. Under the program, schools must do the following: 1) Hire an external evaluator. 2) Perform an assessment. 3) Develop an action plan. Schools that were included in this program achieved notoriety overnight. "They were constantly in the media and the subject of conversation in their communities as a result of their schools' underperforming status. Some have described the media attention as painful." At one elementary school in Fairfield a teacher states in reaction to being accepted into the program, "We had no idea of what we were getting into. Suddenly we were labeled underperforming and always in the newspaper, even though we were working so hard in a very low socioeconomic area. We came up with a T-shirt, 'We are the chosen – II/USB' to make ourselves feel better." According to the article, this program is described as incurring a tremendous amount of pressure on students and teachers. The article concludes, "For some teachers, being in the program was an exciting opportunity to bring about meaningful change. For others, it was traumatizing, degrading, and a waste of time."
23. California	Imperial Valley Press, Heather Bremner (Staff Writer), May 9, 2004).	Teachers, feds fight over law	Story describing the range of teacher reactions to No Child Left Behind. The article notes that for some teachers, the constant emphasis on raising test scores pulls the creativity out of the classroom. According to one Elementary School District Superintendent, California has had a long history with accountability, but the demands of No Child Left Behind changed some things. "Namely, schools were suddenly compared nationally instead of on a statewide scale and by next school year, emergency credentialed teachers will not be accepted. Although Klentschy said NCLB does stress out teachers, he doesn't think it has created a war zone on school grounds. "'I don't think it pits teachers against students,' he [Michael Klentschy, El Centro Elementary School District Superintendent] said. 'I think really what it does is compresses the school day.'" He goes on to note that "with the accountability of these high-stakes tests, I think it does place teachers in a stressful position and that is passed on to students." Diana Harvey, vice president of the Calexico Teachers Association and a Dool Elementary School teacher, said there are numerous elements of NCLB that need mending. She said the tests presume all kids speak English, have the same resources at home, learn at the same pace and enter school at the same level. "I mean, I have some kids who spend vacation on the Greyhound going to the nearest prison and some spend it in Hawaii and I'm expected to give them the same test," Harvey said. Also, she said all the structure doesn't allow her as much creativity. Harvey said there's nothing wrong with teachers being accountable and she has no problem with ensuring teachers are qualified for the classroom. But the programs aren't fully funded and districts are still expected to implement them, she said. Still, she would rather see all the money that's poured into testing into other areas, such as healthier cafeteria food. "We talk about NCLB but does anyone care or look at what we feed them during the day," she asked.

This document is available on the Education Policy Studies Laboratory website at: **http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc**

Table 9, continued

24. Texas	Baytown Sun, Keri Mitchell (Staff Writer) (April 8, 2004).	Survey: Tests hurt teachers' ability to teach	A survey conducted with the Baytown Classroom Teachers Association that polled its 687 members (receiving 66 percent response rate of members – a total of 35 percent of employees from the Goose Creek District). The overwhelming response, according to association president, June Gordon was frustration, "Teachers are frustrated because more and more is expected and demanded of them, and nothing is ever taken away." When asked about how they perceived Goose Creek's administration to treat teachers, more than half, 255, said the district cares only about TAKS scores at all costs. The next highest response from 56 teachers said the district does not care about employees. Also, teachers reported that they lost teaching time, "not only to the TAKS tests but also to required district assessments given to students every six weeks. Campuses spend a minimum of five days every six weeks on the assessments, totaling six weeks in the school year lost to testing."
25. Tennessee	The Commercial Appeal, Aimee Edmonson (Staff) (September 24, 2003).	Teachers slam TCAP – 75 percent polled say test isn't a fair gauge of learning, needs overhaul	Story about many teachers throughout Tennessee lamenting the problems with the statewide TCAP exam that is used to evaluate teacher performance. Teachers are stressed out over being evaluated on the basis of one test score and argue that the test doesn't really show what students can do. Said one teacher: "I have had students who cannot read a word, guess well, and do OK on the test. I have also had extremely bright children totally do awful on the test because of something that happened at home the night before or the morning of the test." Another teacher described her angst at having to repeatedly wake kids up during testing, and of seeing bored children bubbling-in Christmas tree designs on their answer sheets. To top it off, teachers are reeling from the stiff new federal law, No Child Left Behind, which demands higher-than-ever performance for all kids regardless of their background. "We don't want to come across as defensive or trying to place blame," said Kay Williams, Shelby County Education Association president. "It seems like morale is really low right now."
26. Kentucky and national perspective	RESEARCH: Kohn, A. (2000).	*The case against standardized testing: Raising scores, ruining the schools*. Portsmouth, NH: Heinemann	One principal in Kentucky said he watched his colleagues "disappear from the ranks. No one wants to blame it on [high-stakes testing programs], but from my perspective as a practicing principal, many of them made it clear they weren't going to put up with unreasonable demands." (Quoted in Kohn, 2000, p. 27).
27. Maine	Portland Press Herald, Mechele Cooper (Staff) (January 25, 2004).	Rural school systems struggle with special education	Teachers in rural schools are frustrated because of the lack of funds for special education services. Under pressure to comply with No Child Left Behind, monies are tight and small schools in rural areas are not given adequate funds to help special education needs students.

This document is available on the Education Policy Studies Laboratory website at: **http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc**

Table 9, continued

28. National perspective	National Desk, Education Reporter, U.S. Newswire (July 6, 2003).	NEA Annual Meeting Takes Aim at No Child Left Behind Flaws	A report on the National Education Association's position against many of NCLB's mandates. As part of the lobbying effort to try to change the law, "nearly 13,000 e-mails and more than 1,000 phone calls to Congress and state legislators [were made] over six days. They urged support for revisions in the federal elementary and secondary education law – including provisions to ensure that over-reliance on standardized testing not interfere with teaching and learning." Many delegates' stories of how the law left many children behind were shared. For example, a Michigan teacher spoke of an elementary school where art, music, physical activity, and recess all had failed to the budget act. And, hundreds of members shared similar stories.
29. Provo, Utah	The Daily Universe, Neal Legler, (June 24, 2003).	Imposed education reform suffers in Utah	Story about how Utah is having a hard time meeting federal guidelines for placing a "highly qualified" teacher in every classroom. Part of the problem is that sometimes teachers are forced to teach in subjects outside of their expertise. For example, one principal said that it is hard when an art teacher, who only has two periods of art a day, is forced to teach other areas. "In those cases, she has classes that need to be filled but not enough teachers to fill them." Article points out that smaller schools have an even harder time.
30. Pine Level, North Carolina	New York Times, Michael Winerip (October 8, 2003).	How a good school can fail on paper	Principal of an elementary school feels pressure to change how many students are identified as special-needs kids in order to help increase student performance. But, she will not force borderline children to face failure on regular standardized tests. She says, "I couldn't. The borderline children experience enough failure and do not need to be humiliated by a test far beyond their abilities." In this community, the principal knows what students are up against. For example, one student is one of 14 living in the same house with grandmother, and another had to sleep at the bus driver's house the night before the state tests because there was a drug raid going on at home.
31. California	Napa Valley News, Heather Osborn (Staff) (February 8, 2004).	'No Child' leaves some teachers behind	Story about a second year teacher, who in spite of the principal's rave reviews about her, and the many commitments to students she has made in her first year, still isn't deemed "highly qualified" by NCLB. In order to be in compliance with NCLB she had to take standardized tests, during which she became enraged. She notes, "When I was sitting there taking the test, it just made me more enraged … I had already proven myself and now, because of some silly law, I had to take a test to show I am qualified by the new standards." In the Napa area, the articled noted that "new elementary school teachers are being hit the hardest by the rules, but all teachers in Napa must meet the 'highly-qualified' criteria by 2005."
32. National perspective	RESEARCH: Kohn, A. (2000).	*The case against standardized testing: Raising scores, ruining the schools.* Portsmouth, NH: Heinemann	A Florida superintendent noted, "When a low performing child walks into a classroom, instead of being seen as a challenge, or an opportunity for improvement, for the first time since I've been in education, teachers are seeing [him or her] as a liability" (Quoted in Kohn, 2000, p. 28).

This document is available on the Education Policy Studies Laboratory website at: **http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc**

Table 9, continued

33. Pennsylvania	Delco Times, William Bender (March 8, 2004).	In-Depth: Grading our schools. How Delco faired on PSSA Test	Students are being prepped for the statewide high-stakes examination that determines schools' compliance with No Child Left Behind. In preparation for the upcoming test, districts are holding raffles to boost student attendance. Other districts are holding educational pep rallies and running homeroom TV quizzes at others. And the slogans circulating throughout Christopher Columbus Elementary School in Chester are energizing students. According to the principal in Chester, "We're hyping them up more about the test and the competitive nature of doing better than before." At the same time teachers are reassuring students who dread the tests and are prone to freeze up. According to Haverford School District Assistant Superintendent Kathleen Taylor, "It's all based on that one test and boy, the pressure is unbelievable … it's something that is felt by the kids, felt by the teachers and administration, all of us. When you have third and fifth graders really worried, that's not fair." Teachers say the consequences of poor performance put undue pressure on students. "What it ends up being is really most of your time is spent teaching for the test because it's seen as an assessment of the school, the teachers and the district. There's much more we could be looking into," said Joanne DeBoy, principal of Harris Elementary School.

This document is available on the Education Policy Studies Laboratory website at: **http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc**

To illustrate the depth of this problem we borrow again from the book by Jones, Jones, and Hargrove who provide excellent citations, and pertinent quotes on this subject.[92] They report that in one survey almost 90 percent of the teachers agreed that they were under "undue pressure" to improve students' test scores.[93] In two other independent surveys three-quarters of the teachers[94] or two-thirds of the educators[95] reported that their jobs were more stressful since the implementation of high-stakes testing. These educators reported that they experienced such emotions as anger, tension, frustration, anxiety, and depression from the imposition of high-stakes tests. Various teachers explain:

> "A few years ago, I really loved teaching, but this pressure is just so intense…I'm not sure how long I can take it."[96]
>
> "Before [the state program], you know, I could just go with kids if something came up which hooked them. But now if we just start off in a different direction, I get worried we won't get back to what's required, and I have to kind of rein them in. I know they get frustrated, and I sure do. I think, well, is this what I got into teaching for?"[97]
>
> "These tests, and all of this pressure to make kids do well on the tests…it's an insult. It's saying we aren't a profession and we can't be trusted to do our jobs, so high-pressure tactics are necessary to make us behave. They're treating us like stupid children, they're turning us into bad teachers, taking away every bit of pride."[98]
>
> "I have seen some of the most bubbly teachers in this school…in this system and probably in the state, lose their enthusiasm, that zest they had for

This document is available on the Education Policy Studies Laboratory website at:
http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc

teaching [as a result of the testing]. [I have] seen a lot of good teachers leave already, and I'm afraid the numbers are going to become more massive. I think that's when they are going to open their eyes when they see teachers walking out by the droves."[99]

"I'm not the teacher I used to be. I used to be great, and I couldn't wait to get to school every day because I loved being great at what I do. All of the most powerful teaching tools I used to use every day are no good to me now because they don't help children get ready for the test, and it makes me like a robot instead of a teacher."[100]

From the study of teachers reactions to CSAP, the Colorado high-stakes test, we hear these voices:[101]

"We are under so much pressure to get good results there is little time for anything else. CSAP definitely decreases morale."

"I had to cease a lot of projects and other activities and programs in my room to make time for the teaching of test-taking strategies. I felt demoralized and unappreciated by all of the negative press in the newspapers and have doubted myself as an educator for the first time. I'm not sure I would go into this profession if I had to choose all over again. I feel pulled from many directions – to make education more personal and then, from the CSAP – to standardize their learning – forcing them into a box whether they are ready developmentally or not."

"A number of my friends are retiring because they don't want to be associated with a partially proficient school. And from their gut level, they are

This document is available on the Education Policy Studies Laboratory website at:
http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc

giving their all but getting a slap in the face. They know their kids are making advancements."

"I find that it is a demoralizing, stressful situation. Teachers are being judged year to year on completely different students. The pressure put on teachers has increased to the point where teachers will be leaving the profession."

"We can't even get teachers to come into schools like the one I am in because they are afraid that they will be called failing. Why should a young teacher come here when they can go to a wealthy school district? I mean the stigma that the grading has put on all schools in minority neighborhoods is just absolutely incredible. The talk that the good teachers don't come here, it basically frightens away anybody with any ability who doesn't know the community. Why should they come somewhere when they are going to be called a failure? I can't blame those teachers."

These pressures naturally enough have tragic consequences. CNN reports:[102]

Last year Betty Robinson, a married mother of two, and principal of the Simonton school, attended a meeting where Gwinnett County school officials discussed school performance, which is based almost entirely on standardized test scores. Each school is expected to show 5 percent improvement over the previous year. Last year, Simonton, which has a substantial new immigrant population, improved by 4.9 percent.

Simonton had been one of four Gwinnett schools included on a list of 436 Georgia schools that were failing new standards under President Bush's "No Child Left Behind" education plan. Under that program, if Simonton does not improve

This document is available on the Education Policy Studies Laboratory website at:
http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc

enough, students could transfer to other public schools and Simonton would have to pay the busing costs.

Early the next morning, before her students were to take yet another important test, Robinson locked her office door and shot herself in the head.

[It was also reported that] Pamela Relf, a teacher at Middlefield Primary School in Eynesbury, England, drowned herself in 2000, leaving a note behind saying she was upset by her school's performance. A report from the Office for Standards in Education said her school had "serious weaknesses," inadequate leadership and that too many lessons were "unsatisfactory or poor."

"Like her colleagues in this and all schools, there was often the feeling of running to stand still and like all of us, she felt the pressure that resulted," said Brian Rayner, a teacher at Relf's school.

There is little more to say. High-stakes testing is like a disease that saps the strength and vitality out of most, though not all, teachers and administrators. It turns too many of them against their profession and their students. It impoverishes their lives. And if schools are not places where teachers want to be, how can they ever be good places for children to be?

Errors of Scoring and Reporting

In any large-scale testing system, three kinds of errors are possible: errors of *test construction* (test items with more than one correct response), errors of *scoring* (items marked wrong when they are right), and errors of *reporting* (reporting that 50 percent of students passed when actually 75 percent did). Typically, everyone concerned with

This document is available on the Education Policy Studies Laboratory website at:
http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc

testing does everything they can to avoid these types of errors. Nevertheless, errors occur, though in the past the consequences of such errors were not so dire. In today's high-stakes testing environment, however, any such errors can be life-changing, as stories in Table 10 illustrate.

There are several instances where students, teachers, administrators, and even test designers identified test items that were incorrectly written or keyed. In some instances, it is even worse—the questions contain misleading or inaccurate data. For example, in Georgia teachers found significant flaws in the state test's science exam (see Article 2 in Table 10):

> [High school physics teacher] Maley estimated [that] about 10 percent of the questions on the science section, which about 30 percent of students fail each year: Had no best answer because of errors in information provided to students; had multiple correct answers; were ambiguous or were misstatements of science. Department officials acknowledge [that] the acceleration formula and periodic table were wrong, because of a printing error, and two questions were thrown out in scoring because of those mistakes.

A similar situation occurred in Hawaii where testing errors were brought to the attention of officials after teachers, students, and others found errors in the test (see Article 3, Table 10). In the New York Times, Diana Henriques reported:[103]

> During a tutoring session last December, Jennifer Mueller, a high school student in Whitman, Mass., came up with a second correct answer for a question on the state's high school exit exam; an answer that the giant company that designed the test had never anticipated.

This document is available on the Education Policy Studies Laboratory website at:
http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc

> When statewide scores were adjusted to reflect Ms. Mueller's discovery, 95 dejected seniors who had failed the test by one point suddenly found they could graduate after all.

Students in Massachusetts also reported two right answers to an item in the eighth-grade science exam (see Article 8). And in New York, 200,000 fourth graders took a test with an item that might have "confused" the students.

In our own state of Arizona, Professor James Middleton, a mathematics educator and researcher at Arizona State University analyzed AIMS items released in 2001.[104] AIMS is Arizona's high-stakes high school exit exam. Middleton was looking for mathematical accuracy, potential for multiple interpretations (which tends to cause confusion in children in a high-stakes situation, unrelated to their degree of understanding of the content), and realism of the item, in terms of any pragmatic context within which an item might have been embedded.

Of the 38 Core items released, 17 (45 percent) had some problem associated with them that could have caused a consistent measurement error, meaning that the score students received for that item might not have reflected their actual level of understanding of content or skill in problem solving. Of those 17 items, ten had problems significant enough to warrant their removal from the assessment. Here is an example of Middleton's analysis of the test and two of the items from the test designed by CTB/McGraw Hill in 2001:

> The trouble begins on page 1, the AIMS Reference Sheet, on which are placed potentially useful formulas and theorems for the students to use in taking the test. Unfortunately, the students cannot trust the Reference Sheet as the formula for the

This document is available on the Education Policy Studies Laboratory website at:
http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc

Volume of a Sphere is incorrect. Instead of 4/3 pi r^2, the stated formula, the actual formula should be 4/3 pi r^3. Moreover, even if the student caught the mistake, they may not remember the value of pi, since the Key on page one suggests that students use 3.14 or 22/7 as the value for p, [which is] the Greek symbol for rho, not pi.

It gets worse from there…

Problem 16: Alex is building a ramp for a bike competition. He has two rectangular boards. One board is 6 meters long and the other is 5 meters long. If the ramp has to form a right triangle, what should its height be?

A 3 meters

B 4 meters

C 3.3 meters

D 7.8 meters

In this item, none of the answers is correct. The student is expected to use the Pythagorean Theorem ($Hypotenuse^2 = Side1^2 + Side2^2$). So, $(6m)^2 = (5m)^2 + (EF)^2$. To maintain a right triangle, the only correct answer is $(11)^{1/2}$ meters, one that is cumbersome in real life, and so requires rounding off to an acceptable level of accuracy. Depending on the convention for rounding, a reasonable height could be 3 meters (if the convention is rounding to the nearest meter), 3.3 meters (if the convention is rounding to the nearest decimeter), 3.32 meters (if the convention is rounding to the nearest centimeter), and so on.

This document is available on the Education Policy Studies Laboratory website at:
http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc

The answer marked as correct, 3.3 meters is actually about 1.2 centimeters off (about 1/2 inch). Any carpenter worth his or her salt would not make an error of 1/2 inch given a tape measure that is precise to 1/32 inch.

Moreover, as a male, I cringe at the thought of a bike competition that requires riders to jump off 3.3 meter heights (between 10 and 11 feet, ouch!). Or if the rider is to ride down the ramp, a slope of 66 percent (33.5 degrees) is steep enough to scare the bejeebers out of me.

Lastly, a 6 meter board? Come on! When was the last time you found a board of 20 feet at Home Depot? In short, the context within which the problem is embedded shows a lack of the everyday sense for numbers that is required in the elementary standards for Arizona children.

If the released items are a representative sample, then this analysis indicated that over 1/4 of the State's mathematics assessment provide incorrect data to the state department of education, school districts, parents and children anxious to graduate.

But things did not get much better over time. The Arizona State Department of Education just released 18 more items from the spring 2004 administration of the AIMS test. Now only a sixth of these items appear to be wrong or misleading.[105] But if 1 out of 6 items on the AIMS test is mathematically flawed, it could mean that up to 17 percent of a student's responses are marked incorrect when they should not be. For many of Arizona's students, just a few poorly written problems of this sort can cause them to fail the test and not graduate from high school.

This document is available on the Education Policy Studies Laboratory website at:
http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc

Badly constructed tests are common. Here is the Editorial Board of the Orlando Sentinel weighing in:[106]

> As a Sentinel editorial revealed last week, third-graders are likely to be tested on material written at a seventh-grade level or higher. Sentences in the state's sample test were too long-winded and complex for the average young reader.
>
> A look at the state Department of Education's Web site shows that fourth-graders fare no better. The average fourth-grader is likely to encounter FCAT essays that are better suited to much older readers. What's more, one essay contains a spelling error.
>
> An essay about silver ants from a sample fourth-grade test is written at the seventh-grade level, according to the well-known Flesch-Kincaid readability index. The index measures readability based on the length of words and sentences.
>
> Another essay from a previous fourth-grade FCAT also was written at the seventh-grade level, according to the index. The state posts only select items from FCATs given several years ago.
>
> The latter essay, about a butterfly farm, contains challenging words, such as chrysalis and hydrangea and hibiscus. It also refers to gulf "frittilary" butterflies and asks a question about the "frittilaries" in the essay. That word is so tough that it is frequently misspelled -- including on this test. The correct spelling, according to Webster's New World Dictionary and scientific dictionaries and encyclopedias, is "fritillary."

This document is available on the Education Policy Studies Laboratory website at:
http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc

A report by the National Board on Educational Testing and Public Policy has identified many more examples of these and other types of errors—only a few of which we report in Table 10.[107] Releasing a test with so many bad items means that test companies are spending too little on item review panels and on item field testing. But then, their incentive is money, and as is so often the case when money is the major objective, quality suffers. States are also at fault too. States usually have commercial companies bid to get these contracts and usually pick the cheapest of those bids as the contractor. Because of this the state contributes to the process of having too many bad items on too many high-stakes tests.

There are also many examples of scoring errors. As a consequence of being under pressure to get scores back to states, scoring companies oftentimes rush through the scoring process which increases the possibility of making scoring errors. In Minnesota about 8,000 students were unjustly denied a high school diploma because it was determined that they "failed" the test when in fact they did not. Subsequently, many of these students who were "wrongly" failed participated in a suit against the scoring company (NCS Pearson), which they won. The judge in the case was severe, writing "a scathing opinion" that said the company "continually short-staffed the relatively unprofitable Minnesota project…."[108] Compensation did not change the fact that Pearson's placement of profits before quality denied thousands of students the opportunity to participate in the once-in-

Pearson's placement of profits before quality denied thousands of students the opportunity to participate in the once-in-a-lifetime tradition of walking across the stage at graduation. These errors change lives forever.

This document is available on the Education Policy Studies Laboratory website at:
http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc

a-lifetime tradition of walking across the stage at graduation. These errors change lives forever.

The stories of simple human error and avoidable error are almost endless. In New York, 2,400 students in grades 3, 5, 6 and 7 were asked to retake a test (after they were mistakenly given the wrong practice exam), using a test booklet with a key that didn't match the answer booklet. "At one school in Brooklyn, teachers clarified the issue by writing on the blackboard that 'A=E,' 'B=F,' 'C=G,' and 'D=H.'"(Article 20, in Table 10). Were students given extra credit for accurately decoding the adult-created problem with the answer key? Probably not.

There are also examples of *reporting* errors. These kinds of errors have all sorts of consequences. To students, it is likely a humiliating—and confusing—experience to first be told you failed a test only to be later told you passed. How does a student recover? What is she to make of the meaningful of tests in the first place if test "performance" and its consequences can hinge on one item? To teachers and schools, being wrongly labeled "failing" is demoralizing and difficult to recover from as illustrated in Nebraska (Article 9). In Pennsylvania, it didn't happen once, but twice that school's publicly released report cards contained erroneous information, including bad achievement data that affected school level ratings. There are literally hundreds of examples of all three of these types of errors.[109]

The US Government Accounting Office (GAO), as part of its investigation into the functioning of NCLB has looked at this problem too.[110] One of their major findings was that the problems of unreliable tests and test scoring are common. The report notes:

This document is available on the Education Policy Studies Laboratory website at:
http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc

> Concern about the quality and reliability of student data was the most frequently cited impediment to implementing student proficiency requirements….For example, officials in California indicated that they could not obtain racial and ethnic data—used to track the progress of designated student groups—of comparable quality from their school districts. Officials in Illinois reported that about 300 of its 1,055 districts had problems with data accuracy, resulting in those schools' appealing their progress results to the state. Similarly, officials in Indiana acknowledged data problems but said addressing them would be challenging. Inaccurate data may result in states incorrectly identifying schools as not meting annual goals and incorrectly trigger provisions for school choice and supplemental services.

We are not surprised that many of the larger testing companies are involved in many of these cases, though we are sure they did not set out to deliberately and negatively impact the lives of students and their families. But they are the ones that bid low on test development contracts, and then they have to find ways to make a profit. And they apparently do that by sacrificing quality. Harcourt Assessment is attached to several errors we report on. For example, in 1999 in California (Article 4), they were named in a $1.1 million lawsuit for mismanaging the statewide assessment system. What is disturbing is that they, and the states that hire them, don't seem to learn. Later, in 2003, they were also responsible for distributing exams and answer booklets that did not match up in New York (Article 20), and for distributing tests

Everyone is compromised when the integrity of high-stakes decisions must rely upon bureaucrats and budgetary analysts.

This document is available on the Education Policy Studies Laboratory website at:
http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc

that contained errors in Hawaii (Article 3). Other well-known testing companies that have had similar problems include NCS Pearson and CTB/McGraw Hill. Everyone is compromised when the integrity of high-stakes decisions must rely upon bureaucrats and budgetary analysts. Perhaps the problems just discussed are nothing more than poof of the old adage that you get what you pay for.

This document is available on the Education Policy Studies Laboratory website at:
http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc

Table 10: Errors of Scoring and Reporting

Location of Story	Source	Headline	Story
1. National perspective	Chicago Tribune, Stephanie Banchero (Staff Writer) (online edition, November 26, 2003).	Sea of testing data buries U.S. schools: Complex results, errors delaying state report cards	Article decrying how overwhelmed school officials are in trying to meet federal mandates to publicly release school-level achievement information. This pressure, according to the article, has resulted in a barrage of states releasing information prematurely and riddled with errors or delaying the release of information so long that it is useless to parents who may want to use it to make decisions about where to enroll their child. Some examples: Illinois spent $845,000 on a new reporting system, but after problems with data, information was released a month after they were supposed to and even still, data was missing. In Louisiana, hundreds of school report cards were error-ridden after a computer glitch incorrectly indicated whether a group of students had met state standards. In Utah, as of November 26, the state was still trying to get numbers back and off to parents.
2. Georgia	The Atlanta-Journal Constitution, James Salzer (June 3, 2001).	Teachers find flaws in state test's science part	Tim Maley, a North Springs High School physics teacher identified errors in the state's mandatory high school graduation test. Maley noticed a "remarkably high number of errors" on a section of the test that stumped thousands of students, keeping many who couldn't pass it – some who missed by only one answer – from earning a high school diploma. "It just seemed like the test was written with bad science," he said. "It's like they did not understand what they were writing about." Maley estimated about 10 percent of the questions on the science section, which about 30 percent of students fail each year, "had no best answer because of errors in information provided to students," had multiple correct answers, were ambiguous or were misstatements of science. Department officials acknowledge the acceleration formula and periodic table were wrong because of a printing error, and two questions were thrown out in scoring the test because of those mistakes. Some other problems, state staffers said, involved questions that were being "field-tested," items not counted but used to see how high school students answered them. Such questions can be used on future tests.
3. Hawai'i	Honolulu Advertiser, Derrick DePledge (Staff Writer) (May 6, 2004).	Standardized tests checked for errors	Specialists at the Hawaii State Department of Education are combing through standardized tests students took this spring for errors. The potential errors were brought to the attention of officials after test coordinators, teachers, and students spotted numerous mistakes this past spring. The tests were prepared by Harcourt Assessment Inc., a company that has a five-year, $20 million contract with Hawai'i DOE. The state schools superintendent, Pat Hamamoto said that "no student or school would be held to any of the test mistakes." Harcourt has offered to send letters of apology to schools and will likely pay the costs of the inquiry and any remedy.

This document is available on the Education Policy Studies Laboratory website at: http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc

Table 10, continued

4. California	Associated Press, Steve Geissinger (August 3, 1999).	State board penalizes company for errors on school tests	Harcourt Educational Measurement company is stung with a $1.1 million fine because of the company's errors in managing California's 1999 Standardized Testing and reporting program. Harcourt accidentally miscounted about 250,000 students as not fluent in English and erred in the scores for 190,000 students in year-round schools. About 4.2 million children were tested. The severity of mistakes has led to "an unfortunate lack of confidence now in a statewide test that was really meant to send us down a road of not just high stakes but high standards," said board member Monica Lozano.
5. Minneapolis, Minnesota	Star Tribune Duchesne, Paul Drew (July 29, 2000), p. 1A.	8,000 passed test after all	Almost 8,000 high school students who were first told they failed the math section of the state's basic skills test, actually passed (including 336 who were ultimately denied diplomas because of their scores). These students were victims of a scoring error by National Computer Systems (NCS).
6. Minneapolis, Minnesota	Star Tribune, James Walsh (Staff Writer) (February 15, 2003).	5,000 file to claim test-error money: About 2,000 wronged students file	In the last days during which nearly 7,000 eligible students could file for their share of settlement money for being wrongly failed on the statewide test, the company (NCS Pearson) received a barrage of phone calls. The settlement money (upwards of $7 million) is compensating students for tutoring or lost or delayed college careers as the result of a scoring error which led thousands being wrongfully told they'd failed the test. About 50 of them were denied high school diplomas or a walk across the graduation stage.
7. Massachusetts	Lynn Daily Item, (December 11, 2002).	Lynn teachers discover error on MCAS exam	A group of teachers uncovered errors on the MCAS exam meaning more students may have passed it than originally thought. One error was debated by various mathematicians and the original keyed answer stayed even though the math teachers asserted that all four choices could conceivably be correct. The previous week, an additional 449 students passed the exam as a result of one student finding a second answer to one of the tenth-grade math questions.
8. Boston, Massachusetts	Boston Herald.com, Kevin Rothstein (Staff Writer) (December 9, 2003).	2 rights equal better MCAS scores	A discovery of two right answers to an eighth-grade science exam lead to better test scores for 1,367 students. Crediting the second answer meant that 846 eighth graders' scores will go from failing to needs improvement, 447 scores will change from needs improvement to proficient, and 74 go from proficient to advanced.
9. Omaha, Nebraska	Omaha World Herald, Paul Goodsell (May 1, 2002), p. 2B.	State Education Department offers apologies to 7 schools	Nebraska Department of Education placed seven schools on a list of 107 schools needing improvement that should not have been there. They were forced to send out letters apologizing for the miscalculation.

This document is available on the Education Policy Studies Laboratory website at: **http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc**

Table 10, continued

10. Allentown, Pennsylvania	Morning Call, Christina Gostomski (Staff Writer) (November 19, 2003), p. B1.	New schools assessment has errors: Teacher data mistakes will require changes in report released today	For the second time in three months, the State department of Education (under pressure from the federal government) is prematurely releasing a major report with inaccurate data that will require the document to be redone. The 2003 state report card, which details the performance of every school district, contains incorrect information on at least 70 school districts (there are 501 in the state). This report's release follows an earlier released report that contained errors including achievement data (that affected school rankings).
11. Tampa, Florida	Tampa Tribune, Marilyn Brown (Staff Writer) (September 19, 2003).	Florida miscalculates schools' federal marks	Six weeks after the state education department gave most of its public schools the bad new that they didn't measure up to the new federal standards, it was determined that mistakes were made and 60 schools previously determined to have fallen short actually made it.
12. Madison, Wisconsin	Capital Times, Kathryn Kingsbury (Staff Writer) (June 29, 2001).	Schools question scores: Child who didn't take test got results anyway.	Standardized test scores released in June of 2001 (based on an exam published by CTB/McGraw Hill) may have contained errors. The percentage of students rated as proficient or "advanced" in the test subjects dropped significantly this year, which raised some questions about whether the test results were flawed. Further, some parents received test scores for their child even though he/she didn't take the exam. One parent received a score of 430 for her daughter (an "advanced" score was 611).
13. Connecticut	Christina Hall (Staff Writer) (January 30, 2004).	Test scores delayed	Superintendents, principals, teachers, and students are forced to wait an extended amount of time while CTB McGraw-Hill re-scores all of the open-ended items in writing, reading, and mathematics to ensure accurate scores.

This document is available on the Education Policy Studies Laboratory website at: **http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc**

Table 10, continued

14. Michigan	The Herald-Palladium, Kim Strode (Staff Writer) (January 31, 2004).	AYP frustrates school officials	Two reports for measuring schools were released on January 30, 2004 – a statewide ranking, as well as the federal government's AYP ranking. Although MEAP scores were released as raw data in late October, schools couldn't be sure they made AYP until the state officially released the results. Many districts reported inaccuracies or inconsistencies in the state database. South Haven administrators, for example, immediately found errors in reports. South Haven Schools Superintendent Dave Myers said initial data the district received showed Central Elementary did not make AYP, even though state officials earlier had said the school did make AYP after a district appeal. Lawrence Junior High was originally listed in the database as not making AYP. Stoll said the district had been granted an appeal for the school, and it did, in fact, make AYP. Stoll said after a conversation with state officials on Friday, the error in the database was corrected. Paw Paw's Later Elementary School also was awarded an appeal, meaning the school did make AYP. It was not published that way. According to Lori Cross, assistant superintendent for instruction and curriculum, state officials told the district that, because of time constraints, the error may not be corrected on initial releases on the website.
15. Mesa, Arizona	Associated Press (December 1, 2001).	Errors found in AIMS Math scores for three grades	Arizona students in grades three, five and eight received inaccurate math scores on the AIMS test in 2000 and possibly in 2001, according to state officials. The errors came after an announcement in November of 2001 that companies hired by the state calculated inaccurate writing scores on AIMS for third and fifth graders in 2000. According to David Garcia, the state associate superintendent, "the error in writing scores was glaring. The error in math is more technical, but we need to be prudent and get this right."
16. New York	Associated Press (October 22, 1999).	New mistake for company that misreported city school test scores	More than 8,600 New York City school children were sent to summer school or held back because of test score errors by CTB/McGraw Hill.
17. New York	New York Times, Anemona Hartocollis (January 14, 1999).	Skewing of scores is feared in a test for fourth graders	Thousands of children who took the state test's first section had already studied much of the material on which it was based. It was found that as many as one out of every 12 fourth graders in New York used books and a CD-ROM that included the subjects of some of the multiple-choice questions – information that may have invalidated their test scores.

This document is available on the Education Policy Studies Laboratory website at: **http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc**

Table 10, continued

18. New York	Associated Press, Michael Gormley (March 6, 2003).	Multiple choice question in 4th grade test flunks	The state Education Department has omitted a question on a standardized test given to 200,000 fourth graders in February because the answer to the multiple choice question could have confused students. “Once we realized that this might be confusing to some students, we decided to eliminate the question,” state Education Department spokesman Tom Dunn said Thursday. “We wanted to avoid the possibility of any unfairness.” “I think that there is a high degree of sloppiness really approaching incompetence in terms of how these tests are being prepared,” said Ann Cook of the New York Performance Consortium that opposes high-stakes tests. “It makes you seriously question how the state can use these tests to make critically important decisions about these kids' lives.”
19. New York	New York 1 News (May 10, 2004). http://www.inboxrobot.com/	Petition Filed to help settle controversy over Third-grade reading test	A group of parents and lawmakers are going to court to give some students what they call a “fair chance” on the citywide third-grade reading exam that determines whether they will be promoted to the next grade. Allegedly, some students had prepared for the English Language Arts Exam with an old version of the test and about 20 items that appeared on this year's test. As a result, there has been a flurry of controversy over how to handle the problem. DOE suggested that students could either be graded on the 30 questions they had never seen, or they could retake the test. But, many believe this is an unfair solution. According to City Councilwoman Melinda Katz, "This exam, from the beginning, has been riddled with problems,” said Katz. “The DOE are the ones who had the prep material, gave out the prep material, they are the one who sent the prep material home with the parents, and now a few weeks later they are saying to the children, ‘Listen, we made a mistake and now you're going to have to take the exam again because we made the mistake.’ It’s not the right thing to do. The Department of Education should be the first ones to teach that lesson that we all teach our children to accept responsibility for your mistakes.”
20. New York	Queens Chronicle, Robert Brodsky (May 20, 2004).	Testing their patience--citywide third-grade exams riddled with errors	On the heels of the discovery that thousands of students in grades three, five, six, and seven unknowingly studied for the original English Language Arts exam using last year's exam, city officials said that the test questions failed to match the answer key. About 2,400 students were forced to take the make-up exam after finding out they studied from last year's exam for this year's. However, moments into the exam, instructors noticed that the questions did not correspond with the answer booklets. “The test booklets directed students to select from answers A, B, C and D, and on alternating questions, from E, F, G and H. However, the answer documents only offered the students the opportunity to choose from the first four letters. Despite the confusion, administrators continued with the test, instructing students to circle the answers directly on the test booklet. At one school in Brooklyn, teachers clarified the issue by writing on the blackboard that “A=E,” “B=F,” “C=G” and D=H.” Nonetheless, education officials said they do not expect to invalidate the scores. Harcourt Assessment, which published the exams, accepted full responsibility for the mistakes.

This document is available on the Education Policy Studies Laboratory website at: **http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc**

Table 10, continued

21. South Carolina	The Herald, Jennifer Stanley (December 10, 2000).	Parent cautions: PACT no panacea	One Rock Hill parent, Susan Van Zile, "laughed" when she found out that her son had failed the writing portion of the state's standardized test because her son, a sixth grader, was a straight-A student and had won several writing awards. After complaining to the school an investigation was conducted, it was revealed that his "writing deficiency" was a result of a computer error. Given that the state places so much emphasis on test scores for making decisions about students, Susan Van Zile began campaigning against the use of tests as the sole determinant for making decisions about individuals. "Van Zile has continued e-mailing and researching, telling state officials not to use test scores for the school ratings or as a means to decide who may proceed to the next grade and who will be held back. And it's for those same reasons the state Legislature has delayed using test scores as the only basis to promote a child to the next grade."
22. Virginia	The Virginian-Pilot, Alice Warchol & Mathew Bowers (November 4, 2000).	State says education ratings inaccurate by averaging pass rates for past SOLS	The Virginia State Department of Education released school accreditation ratings based upon incomplete calculation, possibly giving dozens of schools statewide lower ratings than warranted. Three elementary schools in Virginia Beach – Kempsville, Pembroke and White Oaks – were incorrectly labeled with the third-lowest of four rankings, "provisionally accredited/needs improvement, state educators confirmed Friday. 'That's three schools that received bad publicity, and they didn't deserve it,' said Kathleen O. Phipps, a Virginia Beach schools spokeswoman."
23. Louisiana	Times-Picayune, Mark Waller (May 17, 2000).	Errors may undermine LEAP Data	Just days after learning which students had failed and which schools had been labeled as failing, the state learned that there were duplications and inconsistencies in the testing data. Combing through the numbers, Jefferson officials have found that some students are listed as attending multiple schools, others are listed in the wrong schools and yet others are listed as failing the test at one school when they actually scored well at another school. The errors represent a problem not only for parents of the failing students who haven't been notified, but also for parents, principals and teachers interested in school rankings, Barton said. "In the past, we've never gotten this information so early," she said. "The intent of the state was to deliver failures, but we also got an alphabetical roster of every child's test scores. While on the surface it would appear that this information would be useful, so far, it's not." State Department of Education spokesman Doug Myers agreed. He said the data released so far is meant to allow school systems to plan their summer LEAP programs for students who failed. In Jefferson, the special summer school begins in two weeks. Numbers released Friday showed that at least 2,500 students, one in three of those who took the test, failed. School officials still are trying to determine an exact number.
24. Minnesota	Pioneer Press, Paul Tosto (March 13, 2004).	Analysis state education: Reliance on testing poses many pitfalls	An article outlining the major pressures felt by educators as a result of increasing stakes attached to testing in the state. State has had difficulties accurately interpreting achievement data resulting in statewide publication of data results that were in error. This year, a huge jump in fifth grade achievement was erroneously reported. The state had a similar problem in 2000 when there was a basic skills math test scoring error by the testing contractor resulting in nearly 8,000 students being told they failed when they actually passed.

This document is available on the Education Policy Studies Laboratory website at: **http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc**

Table 10, continued

25.Georgia	Atlanta-Constitution, Dana Tofig (June 30, 2003).	State fires testing chief for schools	The director of testing for the State Department of Education was fired. Officials declined to comment following the dismissal, but this action follows David J. Harmon's tenure where there were several problems with standardized tests including: delayed and incorrect scores, and where questions used for practice exams appeared on the printed version of several curriculum exams.
26. National perspective	Yahoo news, Ben Feller (AP education writer) (October 1, 2004).	Unreliable Data, Oversight Hampers Ed Law	Unreliable test scores and other shaky data may be causing states to wrongly penalize some schools under federal law, congressional investigators have found. The report is the latest to raise a warning about the accuracy of school data – an essential underpinning of the No Child Left Behind law – among the states. "Measuring achievement with inaccurate data is likely to lead to poor measures of school progress, with education officials and parents making decisions about educational options on the basis of faulty information," said the report by the Government Accountability Office, Congress' auditing arm. Under the law, schools that get federal poverty aid and fail to make enough progress for two straight years must allow any of their students to transfer. If the schools fall short three straight years, students from poor families must be provided a tutor of their choice. But states may be incorrectly triggering the transfer and tutor provisions, the GAO said. Illinois officials reported that about 300 of their 1,055 school districts had problems with data accuracy. California officials said they couldn't get comparable racial and ethnic data across their school districts. Overall, more than half of the state and school district officials interviewed by the GAO said they were hampered by poor and unreliable data.
27. National perspective	Research Report: Rhoades, K., & Madaus, G. (May 2003): http://www.bc.edu/ research/ nbetpp/ statements/ M1N4.pdf	Errors in Standardized Tests: A systemic Problem	A report documenting a series of test errors that have been detected by students, teachers, administrators, and test companies themselves. They report on 52 test errors reported by consumers dating back to 1981 and 26 errors found by testing contractors dating back to 1976. Examples of errors in school rankings are also provided.
	Errors found by Consumers		1. 1981 (PSAT): A Florida student challenged keyed answer to a question about a pyramid. ETS determined the keyed answer was wrong.
			2. 1991 (California Test of Basic Skills): A Superintendent and testing expert noticed that students with similar raw scores were receiving very different local percentile ratings. They noted the errors had existed for at least six years.

This document is available on the Education Policy Studies Laboratory website at: **http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc**

Table 10, continued

			3. 1996 (Stanford 9): Philadelphia district superintendent announced the company that developed the Stanford 9 admitted to scoring errors in 1997 dating back to 1996. The error causes two schools to be classified as needing remedial help when in fact they had improved.
			4. 2000 (Basic Standards Test given in Minnesota): A parent hired an attorney who took two months of sustained effort to see the test his daughter had taken and failed. It turned out that the wrong scoring key was used for the wrong test form. The error caused 7,930 students to be incorrectly informed they had failed.
	Errors detected by the Testing Contractors		1. 1994 (Connecticut Academic Performance Test – CAPT): Harcourt Brace was fined $85,000 for sending out wrong CAPT scores for grades four, six, eight and ten.
			2. 1997-1998. SATII. An error in grading caused scores on math, Japanese reading, and listening tests to be too high, some by as few as 20 points, others by 100.
			3. 2000. Arizona Instrument for Measuring Standards. Some school-wide scores in writing were skewed when 11th graders were mis-identified as 10th graders.
	Errors in school rankings		1. 2000. TAAS. Two schools in Texas claimed data entry errors lowered their school grades.
			2. 2000 (FCAT): Two elementary school ratings were contested in Florida where schools were graded A through F on FCAT scores. Because of state's policy on rounding, schools' ratings ended up lower than expected. This affected schoolwide bonuses. Because school was believed not to have made progress, no one received financial bonuses.
			3. 2001 (CSAP): Many data errors were found on Colorado's state school report cards. Among the mistakes were statistics on students, teachers, and administrators, test scores and some of the school rankings were wrong.
			4. 2002 (Ohio Proficiency Tests): The Department of Education in Ohio mistakenly included 203 out of 415 elementary schools on their list of low performing schools. The problem was blamed on a computer programming error that required schools to submit scores showing increases in both fourth and sixth grade, even though hundreds of schools did not have both of these grades.

This document is available on the Education Policy Studies Laboratory website at: **http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc**

There is one other little noted consequence of pushing commercial publishers into doing tests cheaply. It is under the radar for news reporters, and thus often escapes notice. This problem is related to the "dumbing down" of the test over time. But in this case it is not politics that drives such decisions, as we argued above, it is costs—the costs of making and scoring items that can more closely relate to the constructs in which we have interest. Richard Rothstein describes this problem:[111]

> Consider a typical elementary school reading standard, common in many states, that expects children to be able to identify both the main idea and the supporting details in a passage. There is nothing wrong with such a standard. If state tests actually assessed it, there would be nothing wrong with teachers "teaching to the test." But in actuality, students are more likely to find questions on state tests that simply require identification of details, not the main idea. For example, a passage about Christopher Columbus might ask pupils to identify his ships' names without asking if they understood that, by sailing west, he planned to confirm that the world was spherical. In math, a typical middle-school geometry standard expects students to be able to measure various figures and shapes, like triangles, squares, prisms and cones. Again, that is an appropriate standard, and teachers should prepare students for a test that assessed it. But, in actuality, students are more likely to find questions on state tests that ask only for measurement of the simpler forms, like triangles and squares. It is not unusual to find states claiming that they have "aligned" such tests with their high standards when they have done nothing of the kind.

This document is available on the Education Policy Studies Laboratory website at:
http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc

In this narrowing of vision about what students should know and be able to do as we go from standards to items, we see three things at play. First is the inherent difficulty of getting complex items to be sufficiently reliable for large-scale assessments. This is not at all a simple problem to solve. Second is the managerial problem of getting complex items created and scored in a timely manner. A testing program where new and complex items are required annually is enormously difficult and expensive to manage. And finally, there is the problem of costs for scoring complex and authentic tasks. The latter two problems are managerial, to be solved by increasing both development time and money, should we ever have the will to do so. But that is not likely.

What is most moving about these stories of testing errors and the lack of proper resources for such a serious undertaking as assessment, is the significantly harmful consequences for test takers and their teachers. These harmful effects occur simply as a function of living in high-stakes testing environments. Without the high-stakes being attached to the tests they could be produced more cheaply and quickly, and we could live with less reliable measures of complex phenomena using more authentic assessments. If it was not for the high-stakes attached to the scores on these tests the nation might be tolerant of more error laden and more unreliable tests. Our nation could easily muddle through with not-so-perfect tests and arbitrary cut scores for defining proficiency if the consequences attached to those scores and standards of performance were not so serious. Students, teachers, administrators, parents, and community members have a legal and

If it wasn't for the high-stakes attached to the scores on these tests the nation might be tolerant of more error laden and more unreliable tests.

This document is available on the Education Policy Studies Laboratory website at:
http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc

moral right to the highest quality tests if the scores on those tests are used to make important decisions. If we cannot afford the time and the money to give the public that kind of data, we cannot continue to do high-stakes testing. On the other hand, if such important decisions were not based solely on test scores, and a wider spectrum of performance indicators were used to make decisions about the performance of students and schools, we might get along with much less costly and perhaps even more valid accountability systems.

So how does Campbell's law relate to errors of testing? In these stories we focus more on the problems of test companies rather than those of teachers and students. We find in these stories evidence of the corruption and distortion of responsible business practices by companies trying to maximize profits. The dual concerns of timeliness and cost puts testing companies under great pressure to produce products that are not as good as they should be. The market does not appear to work well in the case of high-stakes test design. When important decisions are to be made about students, teachers and schools, those individuals and institutions have a right to demand high-quality tests and trustworthy data. Despite a cost to taxpayers of tens of millions of dollars annually, the high-stakes testing systems common across the nation do not all meet such standards of quality.

Conclusion

When either failure or thwarted success is contingent on the value of some indicator, we recognize that individuals will feel pressure to influence the indicator so it will prevent that failure or allow for success. This is human, and thus can be codified,

This document is available on the Education Policy Studies Laboratory website at:
http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc

which is what Donald Campbell did when he gave us Campbell's law. The ubiquity of this law surprised us. Campbell posited that any indicator monitoring something anyone believes to be important is a candidate for being influenced, as when the Shell Oil Company inaccurately reported its oil reserves (its stock would fail) or when Houston Independent School District did not accurately report its dropouts (bonuses would be denied). From the often harmless mis-added golf card score, to the extraordinarily dangerous under-reporting of safety violations at nuclear power plants, we found evidence that indicators on which we depend have been compromised.

In education, using newspaper stories, scholarly journal articles, and academic reports and books, we found evidence that high-stakes testing creates environments in which instances of Campbell's law are abundant. It is all too easy to conclude with confidence that there is distortion and corruption of the educational processes for which high-stakes tests are the monitor, and that the indicators we use and the tests themselves have been compromised. We seem to have two choices. For one, we can build an indictor system less subject to the distortions that accompany high-stakes testing. That is possible. The second choice is a continuation of the present trend. But if we continue to monitor public education with high-stakes tests then we know full well that we should expect distortions and corruptions of the measurement system; we will be required to regularly investigate how such distortions and corruption occurs and determine who is responsible; and we

It is all too easy to conclude with confidence that there is distortion and corruption of the educational processes for which high-stakes tests are the monitor, and that the indicators we use and the tests themselves have been compromised.

This document is available on the Education Policy Studies Laboratory website at:
http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc

will also be required to severely punish miscreants when we find them. This will make educators into police, a role they do not desire. The results of our investigation provide abundant evidence that high-stakes testing is a causal factor in many serious problems with which educators and the general citizenry must deal.

In Table 1 we found many instances of cheating by students, teachers, and administrators. This is the most obvious corruption of the indictors and the educators that we have. Given the widespread condemnation of NCLB as an assessment system by many of America's most prestigious scholars, and the 0.8 to 0.9 probability of a school and its teachers being judged a failure because the bill is designed for that to happen, and the possibility of a single test deciding the fate of many of our nations' youth and its teachers, the more remarkable finding is that so few in our educational system are now actually cheating. But given the pressure they are under, that might not be the case for much longer.

In fact, in Table 2, we see evidence that newspapers already suggest that cheating is inevitable. Instead of asking what conditions are responsible for the wave of cheating that we see, most newspapers apparently have accepted cheating as a fact of life. Of course this need not be the case. High-stakes tests are not the only way to evaluate schools and students. It is worth noting that Finland, the highest achieving country in the world in reading, mathematics and science, apparently have *no* standardized tests that resemble ours whatsoever, though they use teacher made tests in their classroom and school accountability system. Their system uses high standards for allowing teachers into the profession, awards high pay and bestows high status to those that enter teaching, provides rigorous and extensive professional development for the teachers, and depends

This document is available on the Education Policy Studies Laboratory website at:
http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc

on trusting relationships to improve academic achievement.[112] Clearly there are highly successful models of how to build a national school system that we should study before assuming that our corrupting high-stakes accountability system is the only one that will work.

In Table 3 we report evidence about gaming the system. Many articles reveal how educators are trying either to win or not to lose in high-stakes testing environments, by manipulating the indicator for their own benefit. In this case they were doing so by pushing the weaker students out, or letting them drop out without helping them stay in school. Thus the test scores of some schools were made to rise artificially. And in this way, the indicator by which we judge the schools' success or failure, has been corrupted—a perfect example of Campbell's law. Worse yet, in our mind, is the revelation that the corruption of the educators was so widespread. The behaviors of those that help push students out or who let students drop out, demonstrates their abandonment of an ethic of caring, the kind of caring for youth that may have brought them into education in the first place. The harsh treatment by educators of special education students, English language learners, and emotionally distressed youth in high-stakes testing environments was also well documented in these stories. For those corrupted by the high-stakes environment, students become merely score suppressors or score increasers—not human beings in their own right. This turned out to be true for the gifted students as well as their less academically talented peers.

For those corrupted by the high-stakes environment, students become merely score suppressors or score increasers—not human beings in their own right.

This document is available on the Education Policy Studies Laboratory website at:
http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc

In Table 4 we found additional evidence of gaming the system through misrepresentation of dropout data, an important indicator by which we judge schools. Other indicators were also compromised, including college entrance rates supplied by high schools, and the passing rates on state tests. We have learned about the ubiquity of Campbell's law and so we should be alerted that when an indicator takes on importance, we must look at it with the most critical eye.

In Table 5 we saw many examples of either cheating or gaming of the evaluation system, depending on whether or not the line between legitimate and illegitimate test preparation has been crossed. But in either case, we saw how too much time was given to test preparation. Such over commitment to preparation limits the time given over to genuine instruction; limits the creativity and resourcefulness of teachers, making them very unhappy; and bores all but the most compliant students.

In Table 6 we saw that the time needed for testing, and the time needed for preparing for the test, along with the narrow range of subject matter areas that are assessed in most high-stakes testing programs combine in an unfortunate way. When such factors occur together they encourage the promotion of a narrow curriculum for our schools. Areas of the curriculum valued by many citizens and educators alike (e.g., creative writing, art, social studies) are being dropped, and a narrow conception of what it means to be competent person is promoted when high-stakes testing is normalized. Both test preparation and a narrowing of the curriculum have the same effect on educational indicators—they distort the construct being measured. For example, the construct that we ordinarily seek to measure is the mathematics knowledge obtained from a rich and varied high school curriculum in an American state. That would inform us if our citizens are

This document is available on the Education Policy Studies Laboratory website at:
http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc

getting the mathematics we think they need for the society in which they will live. But if we directly prepare students for the tests, drilling them on items suspiciously like those on the test, as well as narrowing the curriculum to spend additional time on the subject matter we are testing, then we have distorted the construct we wanted to assess. When we are not clear about the construct we are measuring, we have corrupted the indicator in which we put our faith.

In Table 7 we learned that there was no single indicator system for accurately representing the state of affairs we want to monitor. Different educational indicators highlight different aspects of the system being evaluated. Educators will, of course, reject those indicators that make them look bad, accepting more easily those that make them look good. While this is self-serving, and to be expected, it does point out that no single indicator has the unique capacity of comprehensively assessing the complex world of schooling. It is probable that a multiple indicator system for the formative evaluation of students, schools, teachers, and districts, would be subject to much less corruption than the single high-stakes indicators we currently use for accountability purposes.

In Table 8 we saw how gaming the system works through the (usually) subtle manipulation of passing and failing rates for high-stakes tests. Since no one knows how to draw cut scores sensibly, they are easily subject to corruption and so passing and failing rates can be easily manipulated. Politics, not science, influences these decisions. There is evidence that because of politics and the inevitable corruption of the indicator used in high-stakes testing, more students are passed on the tests in each successive year. This ultimately makes the accountability system purely symbolic, accomplishing little that is of benefit to education.

This document is available on the Education Policy Studies Laboratory website at:
http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc

In Table 9 we saw that the morale of teachers and administrators was seriously and negatively affected by high-stakes testing. Good administrators and good teachers are being driven out of the profession. We are also driving our most sensible teachers out of the schools that serve the poor and out of the grade levels to which are attached high-stakes decisions about their students. High-stakes testing is clearly undermining the morale of the educational work force. Leithwood, Steinback and Jantzi sum up this problem best:[113] They note that historically, the profession of teaching attracted a disproportionate number of people who were extraordinarily dedicated to the mission of children's welfare. Almost every other kind of organization dreams of having a work force that approaches the level of dedication that teachers have for their corporate missions. "Reform minded governments would do well to consider what is lost by squandering such a resource … and what the costs would be of finding an equally effective replacement."[114]

In Table 10 we saw that scoring errors and reporting errors are common, and are exacerbated by the pressures to test frequently and at low cost. The result is that the economic and social lives of students and teachers' are decided by flawed and corrupted indictors.

We close with concerns about high-stakes testing first expressed by Howard Gardner.[115] He asked those who supported high-stakes testing to think of what it means to be educated in a discipline where one learns to think like a scientist, a mathematician, an artist or an historian. He asked them to think of what it meant in the disciplines to pursue the meaning of truth and its equally important opposite, what is false or what is indeterminate. He asked them to think of what it means to understand beauty, and its

This document is available on the Education Policy Studies Laboratory website at:
http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc

equally important opposites, ugliness or kitsch. He challenged them to think of what it means to deal with what is good and what is evil in this world. After we engage in Gardner's exercise, we must ask ourselves if the assessment tools used for high-stakes testing are designed to measure these things, or are they likely to miss them completely?

The scores we get from high-stakes tests cannot be trusted—they are corrupted and distorted. Moreover, such tests cannot adequately measure the important things we really want to measure. Even worse, to us, is the other issue—the people issue. High-stakes testing programs corrupt and distort the people in the educational system and that cannot be good for a profession as vital to our nation as is teaching. We need to stop the wreckage of our public educational system through the use of high-stakes testing as soon as possible.

This document is available on the Education Policy Studies Laboratory website at:
http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc

Notes & References

[1] Campbell, D. T. (1975). Assessing the impact of planned social change. In G. Lyons (Ed.), *Social research and public policies: The Dartmouth/OECD Conference*. (Chapter 1, pp 3-45). Hanover, NH: Dartmouth College, The Public Affairs Center. (p. 35)

[2] This information is cited in:

Favato, P., Mathison, S. & Calalano C. (2003). *A murmur of dissent: A story of resistance to high-stakes testing in New York State*. Paper presented at the meetings of the American Educational Research Association, Chicago, IL.

[3] This information is cited in:

Sutherland, G. (1973). *Policy-making in elementary education 1870-97*. London: Oxford University Press.

[4] Keller, B. (2004, March 23). Denver teachers approve pay-for-performance plan. *Education Week.* Retrieved September 23, 2004 from: http://www.edweek.org/ew/ewstory.cfm?slug=28denver_web.h23

[5] Bracey, G. (2000). *High Stakes Testing* (CERAI-00-32). Tempe, AZ: Arizona State University, College of Education, Education Policy Studies Laboratory. Retrieved November 27, 2004, from: http://www.asu.edu/educ/epsl/EPRU/documents/cerai-00-32.htm

[6] Brennan, R. L. (2004, June). *Revolutions and Evolutions in Current Educational Testing.* Des Moines, Iowa: FINE Foundation and the Iowa Academy of Education, Occasional Research Paper #7. Retrieved November 27, 2004, from: http://www.finefoundation.org/IAE/iae-op-brennan-1.pdf

[7] For more information, see: http://wwwcsteep.bc.edu/

[8] For more information about Fairtest, see:

Neill, M., Guisbond, L., & Schaeffer, B., with Madison, J. & Legeros, L. (2004). *Failing our children. How "No Child Left Behin " undermines quality and equity in education and an accountability model that supports school improvement.* Cambridge, MA: Fairtest. Retrieved December 20, 2004, from: http://www.fairtest.org/Failing%20Our%20Children/Summary%20Report%20-%20final%20color.pdf

[9] Robert Linn has been president of the National Council on Measurement in Education (NCME), the American Educational Research Association (AERA) and is a member of the National Academy of Education (NAE). His views are expressed in:

Linn, R. L. (2004). Assessments and Accountability. *Educational Researcher, 29*(2), pp. 4-14. Retrieved November 27, 2004 at: http://www.aera.net/pubs/er/arts/29-02/linn01.htm

[10] Heubert, J. P. & Hauser, R. M. (Eds.) (1999). *High stakes: Testing for tracking, promotion, and graduation*. Washington, DC: National Academy Press.

[11] Lyle V. Jones is Professor Emeritus of the L.L. Thurstone Psychometric Laboratory at the University of North Carolina in Chapel Hill. His views are expressed in:

Jones, L. (1997). *National tests and education reform: Are they compatible?* Princeton NJ: Educational Testing Service, William H. Angoff Lecture Series. Retrieved November 28, 2004, from: http://www.ets.org/research/pic/jones.html

[12] Kohn, A. (2000). *The case against standardized testing: Raising the scores, ruining the schools.* Portsmouth, NH: Heinemann.

[13] For examples, see: http://www.bc.edu/research/nbetpp/

Clarke, M., Haney, W., & Madaus, G. (2000). *High Stakes Testing and High School Completion*. Boston, MA: Boston College, Lynch School of Education, National Board on Educational Testing and Public Policy.

[14] Ohanian, S. (2002). *What Happened to Recess and Why Are Our Children Struggling in Kindergarten?* New York: McGraw-Hill.

[15] Orfield, G. & Kornhaber, M. L. (Eds.) (2001). *Raising standards or raising barriers? Inequality and high-stakes testing in public education.* New York: The Century Foundation Press.

[16] Stephen Raudenbush is Professor of Education, Statistics and Sociology at the University of Michigan, and is a member of the National Academy of Education. His views are expressed in:

Raudenbush, S. (2004). S*chooling, statistics, and poverty: Can we measure school improvement?* Princeton, NJ: Policy Information Center, Educational Testing Service. Retrieved December 20, 2004, from: http://www.ets.org/research/pic/angoff9.pdf

[17] Campbell, D. T. (1975). Assessing the impact of planned social change. In G. Lyons (Ed.), *Social research and public policies: The Dartmouth/OECD Conference*. (Chapter 1, pp 3-45). Hanover, NH: Dartmouth College, The Public Affairs Center. (p. 35)

[18] Madaus, G. & Clarke, M. (2001). The adverse impact of high-stakes testing on minority students: Evidence from one hundred years of test data. In G. Orfield & M. L. Kornhaber (Eds.). *Raising standards or raising barriers? Inequality and high-stakes testing in public education.* New York: The Century Foundation Press.

[19] The National Council on Measurement in Education (NCME) and the American Psychological Association (APA) joined with American Educational Research Association (AERA) to define the standards for test construction and use. These are available through AERA. See:

American Educational Research Association (1999). *Standards for educational and psychological testing.* Washington, DC: Author.

[20] Baker, G., Gibbons, R., & Murphy, K. J. (1994). Subjective performance measures in optimal incentive contracts. *Quarterly Journal of Economics, 109*, 1125-1156.

[21] Sarche, J. (2004, October) Ex-Qwest exec agrees to guilty plea. *Denver Post online*. Retrieved December 25, 2004, from: http://www.denverpost.com/Stories/0,1413,36~26430~2409085,00.html

[22] Washington Post (2004, September 30). Timeline of Enron's collapse. Author. Retrieved December 25, 2004, from: http://www.washingtonpost.com/wp-dyn/articles/A25624-2002Jan10.html.

Also see:

White, B. & Behr, P. (2003, July 29). Citigroup, J.P. Morgan settle over Enron deals. *Washington Post*, page A01. Retrieved December 25, 2004, from: http://www.washingtonpost.com/ac2/wpdyn?pagename=article&cntentId=A59547-2003Jul28&notFound=true

[23] Douglas, E. (2004, October 22). Edison says safety data were rigged. *Los Angeles Times*. Retrieved October 23, 2004, from: http://us.rd.yahoo.com/dailynews/latimests/ts_latimes/SIG=10po2s8qq/*http://www.latimes.com/_

[24] Johnston, L. (2004, March 3). Ex-Georgia assistant's exam laughable; can you pass? *USA Today*. Retrieved November 28, 2004, from: http://www.usatoday.com/sports/college/mensbasketball/2004-03-03-harrick-exam_x.htm - exam

[25] Baker, G., Gibbons, R., & Murphy, K. J. (1994). Subjective performance measures in optimal incentive contracts. *Quarterly Journal of Economics, 109*, 1125-1156.

Also see:

Kerr, S. (1975, December). On the folly of rewarding A, while hoping for B. *Academy of Management Journal, 18*, 769-783. Retrieved December 25, 2004, from: http://www.geocities.com/Athens/Forum/1650/rewardinga.html

[26] Prendergast, C. (1999). The provisions of incentives in firms. *Journal of Economic Literature, 37*(1), 7-63.

Also see:

Kerr, S. (1975, December). On the folly of rewarding A, while hoping for B. *Academy of Management Journal, 18*, 769-783. Retrieved December 25, 2004, from: http://www.geocities.com/Athens/Forum/1650/rewardinga.html

[27] Skolnick, J. H. (1966). *Justice without trial: Law enforcement in democratic society*. New York: Wiley.

[28] Campbell, D. T. (1975). Assessing the impact of planned social change. In G. Lyons (Ed.), *Social research and public policies: The Dartmouth/OECD Conference*. (Chapter 1, pp 3-45). Hanover, NH: Dartmouth College, The Public Affairs Center.

[29] Berliner, J. S. (1957). *Factory and manager in the U. S. S. R.* Cambridge, MA: Harvard University Press.

[30] Silverstein, K. (2004, July 1). Pentagon alerted to trouble in the ranks. *Los Angeles Times*. Retrieved November 27, 2004, from: http://www.truthout.org/docs_04/070204C.shtml

[31] Wald, M. L. (2003, October 9). Official Says Airport Trainees Knew Questions Before Tests. *New York Times*, p. A 35. Retrieved November 27, 2004, from: http://www.nytimes.com/2003/10/09/national/09SECU.html?ex=1066690834&ei=1&en=a7bd88f3b23b176c

[32] For more information, see: http://www.lexisnexis.com/

This document is available on the Education Policy Studies Laboratory website at:
http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc

[33] For more information, see: http://www.inboxrobot.com/

[34] For more information, see: http://www.google.com/alerts?t=1&hl=en

[35] For more information, see: http://www.edweek.org/

[36] Winerip, M. (2003, October 8). How a good school can fail on paper. *New York Times*, Section B, p. 9

[37] Linn, R. L. (2004). Assessments and Accountability. *Educational Researcher, 29*(2), p. 14. Retrieved November 27, 2004, from: http://www.aera.net/pubs/er/arts/29-02/linn01.htm

[38] Brennan, R. L. (2004, June). Revolutions and evolutions in current educational testing. Des Moines, Iowa: FINE Foundation and the Iowa Academy of Education, Occasional Research Paper #7, p. 2. Retrieved November 27, 2004, from: http://www.finefoundation.org/IAE/iae-op-brennan-1.pdf

[39] Raudenbush, S. (2004). S*chooling, statistics, and poverty: Can we measure school improvement?* Princeton, NJ: Policy Information Center, Educational Testing Service, p. 35. Retrieved December 20, 2004, from: http://www.ets.org/research/pic/angoff9.pdf

[40] *Ibid*.

[41] Jacob, B. A. & Levitt, S. D. (2002). *Rotten apples: An investigation of the prevalence and predictors of teacher cheating*. (Working paper #9413). Cambridge, MA: National Bureau of Economic Research. Available online at: http://www.nber.org/papers/w9413/.

Also see:

Jacob, B. A. & Levitt, S. D. (2002). *Catching cheating teachers: The results of an unusual experiment in implementing theory.* (Working paper #9414). Cambridge, MA: National Bureau of Economic Research. Available online: http://www.nber.org/papers/w9414/

[42] Pedulla, J. J., Abrams, L. M., Madaus, G. F., Russell, M. K., Ramos, M. A., & Miao, J. (2003, March). *Perceived effects of state-mandated testing programs on teaching and learning: Findings from a national survey of teachers*. Boston, MA: Boston College, National Board on Educational Testing and Public Policy.

[43] *Ibid*.

[44] High tech cheating in Korea is reported in:

Seo, H-s & Lee, M-a (2004, December 9). Cheating took place last year. Prosecutors trace new cases at high schools in Gwangju. *JoongAng Daily, Korea*, p. 2.

[45] Nichols, S. & Good, T. (2004). *America's teenagers--myths and realities: Media images, schooling, and the social costs of careless indifference*. Mahwah, NJ: Lawrence Erlbaum and Associates.

[46] Herbert, R. (2003, October 20). Results of pushing out and dropping out. *New York Times*. Retrieved November 15, 2004, from:

http://www.nytimes.com/2003/10/20/opinion/20HERB.html?ex=1067643408&ei=1&en=cbc4483d1129a720

[47] Arenson, K. W. (2004, May 15). More youths opt for G.E.D., skirting high-school hurdle. *New York Times*, Final, Section A, Page 14. Retrieved December 20, 2004, from: http://www.nytimes.com/2004/05/15/education/15GED.html?ex=1085632344&ei=1&en=bfd603967e3cf3f3

[48] Travis, S. (2004, March 15). Delray teen with stress disorder loses bid to avoid taking FCAT. *South Florida Sun-Sentinel.* Retrieved March 20, 2004, from: http://www.sun-sentinel.com/news/local/southflorida/sfl-ptrauma15mar15,0,5435265.story?coll=sfla-home-headlines

[49] Rubin, J. (2004, November 28). Are schools cheating poor learners? Officials say federal rules compel them to focus on pupils more likely to raise test scores. *Los Angeles Times*, p. B1. Retrieved November 21, 2004, from: http://www.latimes.com/news/education/la-me-behind28nov28,1,5804128.story?coll=la-news-learning

[50] For one example, see:

Moon, P. (2004, November 28). AIMS leaves gifted kids in the lurch. Commentary, *East Valley Tribune*, Arizona, p. F3.

[51] For one example, see:

Kossan P. & Melendez, M. (2004, April 22). Students who passed pushed to retake AIMS…to boost district rankings. *Arizona Republic*. Retrieved April 25, 2004, from: http://www.azcentral.com/arizonarepublic/news/articles/0422AIMS22.html#

[52] The pressure to keep every high scoring child is detailed in:

Golden, D. (2004, February 4). Numbers game. With eye on scores, schools fight over gifted kids. Ohio's odd compromise: You can take my student but leave his tests behind. *Wall Street Journal*, pgs. A1, A9.

[53] Schemo, D. J. (2003, July 11). Questions on data cloud luster of Houston schools. *New York Times*. Retrieved July 15, 2003, from: http://www.nytimes.com/2003/07/11/national/11HOUS.html?ex=1058969280&ei=1&en=c74a998f628d9750

[54] Schemo, D. J. (2003, August, 28). For Houston schools, college claims exceed reality. *New York Times*. Retrieved September 2, 2003, from: http://www.nytimes.com/2003/08/28/national/28HOUS.html?ex=1063064177&ei=1&en=729c0aff59ed3a28

[55] *Ibid*.

[56] Data posted June 7, 2004, at: http://www.doe.mass.edu/mcas/2004/results/0604cdprogrpt.pdf

[57] One day after the Department of Education mislead its citizens Wheelock wrote for the Massachusetts Parents Organization:

Wheelock, A. (2004, June 8). *Massachusetts Department of Education "progress report" inflates MCAS "pass rates" for the Class of 2004*. Retrieved December 25, 2004, from: http://www.massparents.org/news/2004/passrate_2004.htm.

An earlier report on the misleading nature of the Departments' calculations had no effect on the way this story was reported. See:

Haney, W., Madaus, G., & Wheelock, A. (2003). *DOE report inflates MCAS pass rates*. Boston, MA: National Board on Educational Testing and Public Policy, Boston College. Retrieved December 25, 2004 from: http://www.bc.edu/research/nbetpp/pdf/doe_press.pdf

[58] Massachusetts Department of Education press release of May 7, 2003. Retrieved December 20, 2004, from: http://www.doe.mass.edu/news/news.asp?id=1503

[59] Haney, W., Madaus, G., & Wheelock, A. (2003). *MCAS pass rate for special ed students inflated by Massachusetts Department of Education; Discounting thousands of students from the count artificially boosts rates and discourages schools from holding on to all students*. Boston MA: National Board on Educational Testing and Public Policy, Boston College. Retrieved December 2, 2004, from: http://www.massparents.org/news/2003/sped_pass_rate.htm

[60] Sacks, P. (2001). *Standardized Minds: The High Price of America's Testing Culture and What We Can Do to Change It*. New York: Perseus Publishing.

[61] McNeil, L. & Valenzuela, A. (2001) The harmful impact of the TAAS System of testing in Texas. In G. Orfield & M. L. Kornhaber (Eds.) (2001). *Raising standards or raising barriers? Inequality and high-stakes testing in public education* (p. 134). New York: The Century Foundation Press.

[62] Mabry, L. (1999). Writing to the rubric: Lingering effects of traditional standardized testing on direct writing assessment. *Phi Delta Kappan, 80*, 673-679.

[63] McNeil, L. & Valenzuela, A. (2001) The harmful impact of the TAAS System of testing in Texas. In G. Orfield & M. L. Kornhaber (Eds.) (2001). *Raising standards or raising barriers? Inequality and high-stakes testing in public education* (p. 133). New York: The Century Foundation Press.

[64] Jones, M. G., Jones, B. D., and Hargrove, T. (2003). *The unintended consequences of high-stakes testing*. Lanham, MD: Rowman & Littlefield.

[65] Kohn, A. (2000). *The case against standardized: Raising the Scores, Ruining the Schools*. Portsmouth, NH: Heinemann. p. 325 (Cited in Jones, Jones and Hargrove, 2003).

[66] Mitchell, K. J. (1997). What happens when school reform and accountability testing meet? *Theory into Practice, 36*(4), pp. 262-265. (Cited in Jones, Jones and Hargrove, 2003).

[67] Calkins, L., Montgomery, K, & Santman, D. (1998). *A teacher's guide to standardized tests: knowledge is power*. Portsmouth, NH: Heineman. p. 2. (Cited in Jones, Jones and Hargrove, 2003).

[68] Jones, M. G., Jones, B. D., Hardin, B. H., Chapman, L., Yarbrough, T., & Davis, M. (1999). The impact of high-stakes testing on teachers and students in North Carolina. *Phi Delta Kappan, 81*, pp. 199-203.

[69] Smith, M. L. (1991). Meanings of test preparation. *American Educational Research Journal*, *28*(3), pp. 521-542.

[70] Perreault, G. (2000). The classroom impact of high-stress testing. *Education*, No. 4, pp. 705-710. (Cited in Jones, Jones and Hargrove, 2003).

[71] Jones, B. D. & Egley, R. J. (2004, August 9). Voices from the Frontlines: Teachers' Perceptions of High-Stakes Testing. *Education Policy Analysis Archives*, *12*(39). Retrieved December 2, 2004, from: http://epaa.asu.edu/epaa/v12n39/

[72] Taylor, G., Shepard, L., Kinner, F., & Rosenthal, J. (2003, February). A survey of teachers' perspectives on high-stakes testing in Colorado: What gets taught, what gets lost. (CSE Technical Report 588: CRESST/CREDE/University of Colorado at Boulder). Los Angeles: University of California.

[73] Wenglinsky, H. (2004, November 23). Closing the racial achievement gap: The role of reforming instructional practices. *Education Policy Analysis Archives*, 12(64). Retrieved November 27, 2004, from: http://epaa.asu.edu/epaa/v12n64/

[74] Trejos, N. (2004, March 14). Time may be up for naps in pre-k classes. *Washington Post*, p. A01. Retrieved November 27, 2004, from: http://www.washingtonpost.com/wp-dyn/articles/A58706-2004Mar14.html

[75] Waltham Daily News Tribune (2004, February 27). Class requirements cut into recess time. Retrieved March 15, 2004, from: http://www.dailynewstribune.com/news/local_regional/walt_recess02272003.htm

[76] Aquilar, A. (2004, September 28). Many Illinois districts opt out of PE requirement. *St. Louis Post-Dispatch.*

[77] Seattle Post-Intelligencer (2004, September 21). Leave time for recess. *Author*. Retrieved December 25, 2004, from: http://seattletimes.nwsource.com/html/editorialsopinion/2002120763_recessed16.html

[78] Personal communication from a teacher who wishes to remain anonymous.

[79] Zastrow, C. & Janc, H. (2004, March). Academic atrophy: The condition of the liberal arts in America's public Schools. A report to the Carnegie Corporation of New York. Washington, DC: Council for Basic Education.

[80] Taylor, G., Shepard, L., Kinner, F., & Rosenthal, J. (2003, February). *A survey of teachers' perspectives on high-stakes testing in Colorado: What gets taught, what gets lost.* Center for the Study Evaluation Technical Report 588. Los Angeles: University of California at Los Angeles, Graduate School of Education, National Center for Research on Evaluation, Standards, and Student Testing (CRESST) and Boulder, CO: University of Colorado Center for Research on Education, Diversity & Excellence (CREDE).

[81] In Jones, B. D. & Egley, R. J. (2004, August 9). Voices from the Frontlines: Teachers' Perceptions of High-Stakes Testing. *Education Policy Analysis Archives*, *12*(39). Retrieved December 2, 2004, from: http://epaa.asu.edu/epaa/v12n39/

This document is available on the Education Policy Studies Laboratory website at:
http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc

[82] Jones, M. G., Jones, B. D., and Hargrove, T. (2003). *The unintended consequences of high-stakes testing*. Lanham, MD: Rowman & Littlefield. p. 37.

[83] Berman, M. (2004, March 9). Doomed schools: No matter what they do, they'll only continue to fail with a flawed ratings system. Commentary. *The Albuquerque Tribune*.

[84] Kane, T. J. & Staiger, D. O. (2002). *Volitility in school test scores: Implications for test-based accountability systems*. Brookings Papers on Education Policy. Washington, DC: The Brookings Institution.

[85] Glass, Gene V (1978). Standards and criteria. *Journal of Educational Measurement, 15*, pp. 237-261. Retrieved December 3, 2004, from: http://glass.ed.asu.edu/gene/papers/standards/.

Also see:

Horn, C., Ramos, M., Blumer, I., & Madaus, G. (2000, April). Cut scores: results may vary. Boston College, Monographs of the National Board on Educational Testing and Public Policy, *1*(1).

[86] Glass, Gene V (1978). Standards and criteria. *Journal of Educational Measurement, 15*, pp. 237-261. Retrieved December 3, 2004, from: http://glass.ed.asu.edu/gene/papers/standards/.

[87] J. S., Kirby, S. N., Barney, H., Darilek, H., & Magee, S. J. (2004). *Achieving state and national literacy goals, a long uphill road*. A Report to Carnegie Corporation of New York. Santa Monica, CA: RAND Corporation.

[88] Pellegrino, J. W., Jones, L. R., & Mitchell, K. J. (Eds.). (1998). *Grading the nation's report card: evaluating NAEP and transforming the assessment of educational progress*. Committee on the Evaluation of National Assessments of Educational Progress, Board on Testing and Assessment, Commission on Behavioral and Social Sciences and Education, National Research Council. Washington, DC: National Academy Press. p. 182.

[89] Ganesh, A. (2004). Bubbling in: Standardized testing in second grade. Unpublished doctoral dissertation, College of Education, Arizona State University.

[90] Harlow, A. (2003, July). *Why students answer TIMSS science test items the way they do*. Paper presented at the Australian Science Education Research Association, Melbourne, Australia.

[91] Shavelson, R., Baxter, G., & Pine, J. (1992). Performance assessments: Political rhetoric and measurement reality. *Educational Researcher, 21*(4), pp. 22-27.

[92] Jones, M. G., Jones, B. D., & Hargrove, T. (2003). The unintended consequences of high-stakes testing. Lanham, MD: Rowman & Littlefield.

[93] Koretz, D. Mitchell, K., Barron, S., & Keith, S. (1996). *Final Report: Perceived effects of the Maryland school performance assessment program*. Center for the Study Evaluation Technical report 409. Los Angeles: University of California at Los Angeles, Graduate School of Education, National Center for Research on Evaluation, Standards, and Student Testing (CRESST). (Cited in Jones, Jones and Hargrove, 2003).

This document is available on the Education Policy Studies Laboratory website at:
http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc

[94] Jones, M. G., Jones, B. D., Hardin, B. H., Chapman, L., Yarbrough, T., & Davis, M. (1999). The impact of high-stakes testing on teachers and students in North Carolina. *Phi Delta Kappan, 81*, pp. 199-203. (Cited in Jones, Jones and Hargrove, 2003).

[95] Elliot, J. (2000). *NCAE issues recommendations to improve ABCs accountability plan*. Retrieved December 25, 2004, from: http://www.ncae.org/news/000717/pr.shml

[96] Barksdale-Ladd, M. A. & Thomas, K. F. (2000). What's at stake in high-stakes testing: Teachers and parents speak out. *Journal of Teacher Education, 51*, p. 390. (Cited in Jones, Jones and Hargrove, 2003).

[97] Perreault, G. (2000). The classroom impact of high-stress testing. *Education*, No. 4, p. 707. (Cited in Jones, Jones and Hargrove, 2003).

[98] Barksdale-Ladd, M. A. & Thomas, K. F. (2000). What's at stake in high-stakes testing: Teachers and parents speak out. *Journal of Teacher Education, 51*, p. 392. (Cited in Jones, Jones and Hargrove, 2003).

[99] Yarbrough, T. L. (1999). Teacher perceptions of the North Carolina ABC program and the relationship to classroom practice. Unpublished doctoral dissertation, University of North Carolina at Chapel Hill, p. 86.

[100] Barksdale-Ladd, M. A. & Thomas, K. F. (2000). What's at stake in high-stakes testing: Teachers and parents speak out. *Journal of Teacher Education, 51*, p. 392. (Cited in Jones, Jones and Hargrove, 2003).

[101] Taylor, G., Shepard, L., Kinner, F., & Rosenthal, J. (2003, February). *A survey of teachers' perspectives on high-stakes testing in Colorado: What gets taught, what gets lost*. Center for the Study Evaluation Technical Report 588. Los Angeles: University of California at Los Angeles, Graduate School of Education, National Center for Research on Evaluation, Standards, and Student Testing (CRESST) and Boulder, CO: University of Colorado Center for Research on Education, Diversity & Excellence (CREDE). Pgs. 26, 27, and 49.

[102] Cnn.com/education (2002, May 7). Teachers feel stress of high-stakes testing: Results can make or break some educators' careers. Retrieved November 27, 2004, from: http://cnnstudentnews.cnn.com/2002/fyi/teachers.ednews/05/07/testing.pressure.ap/

[103] Henriques, D. B. (2003, September 3). Rising Demands for Testing Push Limits of Its Accuracy. *New York Times*. Retrieved December 25, 2004, from: http://www.cta.org/TestingFacts/TestingPushLimits.htm

[104] Middleton, J. A. (2001). Analysis of the mathematics portion of the Arizona Instrument to Measure Standards (AIMS) High School Form A Released Items. *Electronic Collaborative for Excellence in the Preparation of Teachers*. Retrieved December 25, 2004, from: http://ecept.net/math/2150214.html

[105] Middleton, J. A. (2004). *Critique of released items on the Arizona Instrument to Measure Standards—Mathematics Portion of the 2004 Administration*. Tempe, AZ: Arizona State University, College of Education. Unpublished Manuscript.

This document is available on the Education Policy Studies Laboratory website at:
http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc

[106] Orlando Sentinel (2003, March 19). FCAT failures. Editorial. *Author*. Retrieved April 20, 2003, from: http://www.orlandosentinel.com/news/opinion/orl-edped191031903mar19,0,4532284.story?coll=orl%2Dopinion%2Dheadlines

[107] The full report is by Rhoades, K., & Madaus, G. (2003, May). *Errors in standardized tests. A systemic problem*. Boston, MA: Boston College. National Board on Educational Testing and Public Policy. Retrieved September 21, 2003, from: http://www.bc.edu/research/nbetpp/statements/m1n4.pdf

[108] Henriques, D. B. (2003, September 3). Rising Demands for Testing Push Limits of Its Accuracy. *New York Times*. Retrieved December 25, 2004, from: http://www.cta.org/TestingFacts/TestingPushLimits.htm

[109] Rhoades, K. & Madaus, G. (2003, May). *Errors in standardized tests. A systemic problem*. Boston, MA: Boston College. National Board on Educational Testing and Public Policy. Retrieved September 21, 2003, from: http://www.bc.edu/research/nbetpp/statements/m1n4.pdf

[110] Government Accountability Office (2004, September). *No Child Left Behind Act: Improvements needed in education's process for tracking states' implementation of key provisions* (GAO-04-734). Washington, DC: United States Government Accountability Office.

[111] Rothstein, R. (2004). Testing Our Patience: Standardized tests have their uses. But current federal law uses testing to destroy learning. *The American Prospect*, *15*(2). Retrieved December 26, 2004, from: http://www.prospect.org/print/V15/2/rothstein-r.html

[112] Valijarvi, J. (2004, December). *Assuring a quality education for all: How the Finnish educational system responds to the challenge of equity*. Paper given at the 3rd International Symposium on Educational Policy and Reform in a Global Age December 11-12, College of Education, University of Tokyo, Tokyo, Japan.

Also see:

Lemke, M., Sen, A., Pahlke, E., Partelow, L., Miller, D., Williams, T., Kastberg, D., & Jocelyn, L. (2004) *International Outcomes of Learning in Mathematics Literacy and Problem Solving: PISA 2003 Results from the U.S. Perspective*. Washington DC: Department of Education, National Center for Educational Statistics Report NCES 2005-003. Retrieved December 26, 2004, from: http://nces.ed.gov/surveys/pisa/PISA2003Highlights.asp?Quest=2.

[113] Leithwood, K., Steinbach, R., Jantzi, D. (2002). School leadership and teachers' motivation to implement accountability policies. *Educational Administration Quarterly, 38*, 1, pp. 94-119.

[114] Ibid, p. 115.

[115] Gardner has been a strong critic of testing that fails to represent the important ideas derived from the disciplines. His thoughts about the power of the disciplines to illuminate ideas like truth, beauty, and goodness are expressed in:

Gardner, H. (1999) *The Disciplined Mind: Beyond Facts And Standardized Tests, The K-12 Education That Every Child Deserves*. New York: Simon and Schuster.

This document is available on the Education Policy Studies Laboratory website at:
http://www.asu.edu/educ/epsl/EPRU/documents/EPSL-0503-101-EPRU.doc