DOCUMENT RESUME

ED 301 580 TM 012 440

AUTHOR Medina, Noe; Neill, D. Monty
TITLE Fallout from the Testing Explosion: How 100 Million Standardized Exams Undermine Equity and Excellence in America's Public Schools.
INSTITUTION National Center for Fair and Open Testing (FairTest), Cambridge, MA.
PUB DATE Jun 88
NOTE 50p.
AVAILABLE FROM FairTest, P.O. Box 1272, Harvard Square Station, Cambridge, MA 02238 ($8.95).
PUB TYPE Reports - Evaluative/Feasibility (142)

EDRS PRICE MF01 Plus Postage. PC Not Available from EDRS.
DESCRIPTORS Achievement Tests; Administrators; Basic Skills; *Educational Quality; Elementary Secondary Education; *Equal Education; Minimum Competency Testing; *National Surveys; *Public Schools; School Districts; *Standardized Tests; *State Programs; Test Bias; Test Construction; Testing Programs; Test Results
IDENTIFIERS Telephone Interviews

ABSTRACT

Based on telephone interviews with educational officials from all 50 states and the District of Columbia, an overview of the use and impact of standardized tests in the United States during the 1986-87 school year was developed. In addition to interviewing officials from all state departments of education, officials from 56 sample school districts in 38 states were questioned. This FairTest survey focused on the use of standardized achievement, competency, and basic skills tests. Additional information was gathered by examining recent surveys documenting the use of standardized exams, including intelligence tests, behavioral tests, readiness tests for young children, and placement tests. During 1986-87, the sampled schools administered more than 7.8 million standardized tests to over 5.7 million students to meet local testing mandates--a rate of 1.33 tests per student. Results indicate that: (1) the number of states mandating testing has greatly increased in recent years; (2) southern states and larger school districts test more often; (3) flaws in construction, validation, administration, and use undermine claims of objectivity and produce test results that are inaccurate, unreliable, or biased; (4) racial, ethnic, and sex bias along with bias against groups of low socioeconomic status are the most prominent forms of test bias; (5) American public school have begun to treat standardized tests as the all-purpose answer for promoting school improvement and insuring accountability; and (6) tests heavily influence curriculum development. An agenda for test reform is outlined. An overview of the study methodology, an annotated bibliography, and a press release are appended. (TJH)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

9879

ED301580

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy

"PERMISSION TO REPRODUCE THIS MATERIAL IN MICROFICHE ONLY HAS BEEN GRANTED BY

Monty Neill

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Fallout From the Testing Explosion:

How 100 Million Standardized Exams Undermine Equity and Excellence in America's Public Schools

by Noe Medina and D. Monty Neill

with the staff of the
National Center for Fair and Open Testing

FairTest
National Center for Fair & Open Testing

Box 1272, Harvard Square
Cambridge, MA 02238
(617) 864-4810

June 1988

A National Center for Fair and Open Testing Report

Fallout From the Testing Explosion:

How 100 Million Standardized Exams Undermine Equity and Excellence in America's Public Schools

by Noe Medina and D. Monty Neill

with the staff of the
National Center for Fair and Open Testing

FairTest Staff

Mike Friedman
Noe Medina
Nancy Murray
Monty Neill
Bob Schaeffer
Sarah Stockwell
John Weiss

FairTest
National Center for Fair & Open Testing

Box 1272, Harvard Square
Cambridge, MA 02238
(617) 864-4810

© National Center for Fair and Open Testing (FairTest), June 1988

ERIC

Contents

FALLOUT FROM THE TESTING EXPLOSION:

HOW 100 MILLION STANDARDIZED EXAMS UNDERMINE EQUITY AND EXCELLENCE IN AMERICA'S PUBLIC SCHOOLS

> "Standardized testing seems to have become the coin of the educational realm. In recent years, it seems that the aims of education and the business of our schools are addressed not so much in terms of curriculum . . . as in terms of what gets tested."[1]
>
> - George Madaus & Walt Haney

Standardized tests dominate the educational landscape in contemporary America. Based on a recent study, the National Center for Fair & Open Testing (FairTest) conservatively estimates that public schools in the United States administered over 100 million standardized tests to their 39.8 million students during the 1986-87 school year — an average of more than two and one-half tests per student per year.

Public schools administered over 100 million standardized tests during the 1986-87 school year, an average of more than two and one-half per student per year.

Standardized test results have become the major criteria for a wide range of school decisions. Test scores limit the programs students enter or dictate the ones in which they are placed. Some tests decide who will be promoted and who will be retained in grade; others determine which students graduate from high school. Test results are used to assess the quality of teachers, administrators, schools and whole school systems.

Test proponents, of course, applaud these trends in the public schools. They see tests as "objective" mechanisms to inject accountability into the schools and thereby improve student achievement, staff competence and educational quality. They see standardized exams as essential elements of the "School Reform Movement."

In fact, experience with standardized test use paints quite a different picture. Rather than being "objective" instruments, standardized tests often produce results that are *inaccurate, inconsistent and biased* against minority, female and low-income students. Rather than promoting accountability, tests *shift control and authority* into the hands of an unregulated testing industry. By narrowing the curriculum, frustrating teachers, and driving students out of school, tests *undermine school improvement* rather than advance its cause.

Standardized tests often produce results that are inaccurate, inconsistent and biased against minority, female and low-income students.

Relying on standardized tests will lead to a weaker, not stronger, educational system.

FairTest concurs with the National Academy of Education that "the nation has a right to know what students achieve, what schools are doing, and what more should be done."[2] Standardized tests, when properly constructed, validated, administered and used, can play a role in this effort. Too often however, standardized tests are not properly constructed, validated, administered and used. As a result, relying on standardized tests as the primary criterion for making various school decisions will lead to a *worse, not better*, public understanding of the schools and a *weaker, not a stronger*, educational system.

I. TEST USE IN U.S. SCHOOLS

During the 1986-87 school year, American educators reported that at least 94 to 105 million standardized tests were administered to 39.8 million elementary and secondary public school students. This includes:

* 38.9 million standardized achievement, competency and basic skills tests administered to fulfill local testing mandates;

* 17.5 million standardized achievement, competency and basic skills tests administered in 42 states and the District of Columbia to fulfill state testing mandates;

* between 30 and 40 million standardized tests administered to compensatory and special education students;[3]

* between 1.5 and 1.75 million screening tests for kindergarten and pre-kindergarten students; and[4]

* between 6 and 7 million college and secondary school admissions, Graduate Equivalency Degree (GED) and National Assessment of Educational Progress (NAEP) tests.

This data was gathered by FairTest staff through a series of telephone interviews with officials from all 50 state departments of education, from the District of Columbia school district and from 56 sample school districts in 38 states. Additional information was gathered by examining recent surveys documenting the use of other standardized exams, including IQ tests, behavioral tests, readiness tests for young children and placement tests [see Appendix].

FairTest interviewed education officials from all 50 states and the District of Columbia.

This estimate of 94-105 million tests is probably a conservative one. The total does not include tests administered to identify or place gifted or limited-English proficient students (for which there are no reliable figures). Nor does it include tests administered by private and parochial schools to their students.

Test use may also have been underreported by some school officials. For example, interviews with sources outside the New York City school system uncovered the use of a number of tests which had not been reported by school officials in the FairTest survey.

The survey did reveal that the number of states which mandate testing has greatly increased in recent years. Compared to the findings of a 50-state survey conducted by Education Week in 1985:

The number of states which mandate testing has greatly increased in recent years.

* the number of states requiring students to pass a standardized test for high school graduation increased from 15 in 1985 to 24 in 1987;

* the number of states employing standardized tests to determine whether students should be promoted to the next grade increased from 8 in 1985 to 12 in 1987; and

* the number of states using standardized tests as part of a state assessment program increased from 37 in 1985 to 42 in 1987.[5]

The FairTest survey also revealed three significant patterns of standardized test use in public schools. First, eleven Southern states (Alabama, Arkansas, Florida, Georgia, Kentucky, Louisiana, Mississippi, North Carolina, South Carolina, Tennessee and Virginia) administered standardized tests to fulfill state testing mandates at a rate more than twice that of schools in the remainder of the Nation. In fact, the five states mandating the most tests per pupil (Virginia, Kentucky, North Carolina, Alabama, and Georgia) were all located in the South.

Second, seven states (Iowa, Minnesota, Montana, North Dakota, Ohio, Vermont, and Wyoming) which do not have a state testing mandate all have relatively small minority enrollments in their public schools (ranging from a low of 1% in Vermont to a high of only 15% in Ohio, compared to a national average of about 25%). Alaska is the only state without a testing mandate which also has a substantial minority student population (primarily Native Americans). However, Alaska plans to institute a state testing program in the 1988-89 school year.

Southern states and larger school districts test more often.

Finally, schools in larger school districts are likely to administer standardized tests to fulfill local mandates at a higher rate than schools in smaller districts. The largest school districts (those with student enrollments exceeding 100,000) administered tests at a rate one and one-quarter times that of medium-sized districts (those with enrollments between 25,000 and 100,000) and one and one-half times that of the smallest districts (those with enrollments less than 25,000). However, despite these overall trends, the rate of test administration did vary considerably among the districts.

II. PROBLEMS WITH STANDARDIZED TESTS

"The importance of understanding what it is that tests can and cannot tell us is critical. Not all tests are accurate measures of the skills and knowledge they purport to measure and even the more accurate tests are at best approximations."[6]

- Congressional Budget Office

Standardized tests are consistently sold as scientifically developed instruments which simply, objectively and reliably measure student achievement, abilities or skills. In reality, there are serious problems in the construction, validation, administration and use of standardized tests and their results.

Standardized tests are constructed in ways that often guarantee biased results against minorities, females and low-income students. Test results are evaluated and scored in ways that are often at odds with modern theories of intelligence and child development. The test validation process is often inadequate and far from objective. Many tests are administered in an environment that undermines any claims they may have to being "standardized". Even those that adhere to "standard" administration practices may be biasing the results against minorities, low-income students and females by using examiners who are unfamiliar to the test- takers and by using timed tests.

Flaws in construction, validation, administration and use undermine claims of objectivity and produce test results that are inaccurate, unreliable or biased.

These flaws undermine testmakers' claims of objectivity and produce test results that are inaccurate, unreliable or biased. Ultimately, many tests fail to effectively measure test-takers' achievement, abilities or skills.

Test Bias

Because most standardized tests are written by and for the middle- to upper-class White population, their results often fail to accurately measure the performance of those who do not fit this category. Testmakers claim that the lower test scores of racial and ethnic minorities and of low income students simply reflect the bias and inequity that exists in American schools and society. While such biases and inequities certainly exist, standardized tests do not just reflect their impact, they compound them.

Joseph Gannon provided documentation for this conclusion in a 1980 study for the National Conference of Black Lawyers. Gannon examined the difference in LSAT scores of minority and White college seniors in the same universities and with comparable undergraduate grade point averages. Even after controlling for these characteristics

a gap of 100 points remained when Black and Hispanic scores were compared with those of White students even though they had demonstrated equal academic ability in college.[7]

Researchers have identified several characteristics of standardized tests which could bias results against minorities and low income students. Each characteristic reflects the middle- to upper-class White focus of such tests. As such, test results are as much a measure of race/ethnicity or income as they are of achievement, ability or skill.

Researchers have identified several characteristics of standardized tests which could bias results against minorities and low income students.

Some of these characteristics could also lead to gender bias in standardized tests. However, gender bias affects both males and females. One finds that among very young children, some tests appear to be biased against boys. On the other hand, among older children and adolescents, most bias affects girls.

Several of the characteristics that bias test results relate to language. To communicate their level of achievement, ability or skills, test-takers must understand the language of the test. Obviously, tests written in English cannot effectively assess the achievement, skills or abilities of students who primarily speak Spanish or some other language.

Many groups of English-speakers are affected by a similar, but more subtle, form of language bias. Most standardized tests are written in an elaborated, stylized language rather than the simple and common vernacular. Researchers have discovered that the use of such forms of English prevents tests from accurately measuring achievement, ability or skills of students who use nonstandard English dialects. This includes speakers of Black and Hispanic dialects and of White Southern, Appalachian, and working-class dialects.[8]

A related type of bias stems from stylistic or interpretive differences related to culture, income or gender. For instance, the word "environment" is often associated by Black students with terms such as "home" or "people". White students tend to associate it with "air", "clean" or "earth". Neither usage is wrong; one simply centers on the social environment while the other centers on the natural one. Unfortunately, on a standardized test only one of these two usages, generally the one reflecting the White perspective, will be acceptable.[9]

Standardized tests assume that all individuals perceive information and solve problems in the same way.

Similarly, researchers have discovered that individuals exhibit "different ways of knowing and problem-solving" which reflect different styles, not different abilities. These differences are often correlated

with race/ethnicity, income level and gender. Yet standardized tests assume that all individuals perceive information and solve problems in the same way. Again, assumptions about the universal application of a style exhibited primarily by middle- to upper-class White males limits the reliability of test results.[10]

A third cause of test bias is apparent in questions which assume a cultural experience and perspective which not all children share. "Correct" answers to such questions usually reflect the experiences and perspectives of children and adults from a White middle- to upper-class background. Answers which draw upon the often different experiences and perspectives of racial/ethnic minorities, children from poor families, and children from inner city or rural backgrounds are ignored. Although they may be correct in these different geographical or cultural contexts, they are generally counted "incorrect" on the test.

Research on Mexican-Americans, Blacks and girls all reveal that "items with content reference of special interest" to each group seem to improve their test scores.

The WISC-R IQ test, for example, asks "What is the thing to do when you cut your finger?" The best response, according to the test, is to "put a bandaid on it." Partial credit is also given for a response of "go to the hospital." No credit is awarded for responses of "cry, bleed or suck on it." "Minority children usually perform poorly on this item. A few years ago a Baltimore, Maryland sociologist asked several inner-city youths why they answered the question the way they did. She found that many of these kids answered 'go to the hospital' because they thought that cut meant a big cut. Or they thought it was a small cut — and since they didn't have any bandaids at home, they answered 'suck on it' — and received no points." [11]

Finally, students tend to perform better on tests when they identify with the subjects of the test questions. Research on Mexican-Americans, Blacks and girls all reveal that "items with content reference of special interest" to each group seem to improve their test scores.[12] Unfortunately, standardized tests remain dominated by questions about and for White middle- to upper-class males.[13]

Test Construction

In constructing a standardized test, test developers make a number of unwarranted assumptions. Among the most important are assumptions that results can be sorted so as to fit on a linear scale, and that they can be reported in the form of a single score.

Underlying the former assumption is the belief that development of the knowledge, abilities or skills being tested occurs in a relatively

Thus standardized tests mislabel many, if not most, individuals who exhibit developmental pa terns that differ from a defined "norm".

consistent fashion among all individuals. In fact, developmental researchers tell us that this simply is not true. Child language research, for example, demonstrates that "some children develop the use of pronouns before the development of an extensive noun vocabulary. For others, the reverse pattern of development occurs." Neither is considered to reflect a learning disorder or disability. They simply reflect variations in development patterns.[14] Thus standardized tests mislabel many, if not most, individuals who exhibit developmental patterns that differ from a defined "norm" (based on majority group practice) as being delayed or disordered in their development.[15]

The use of a linear scale not only creates false differences, it masks real differences. Assume, for example, that one student can compute using addition, subtraction, multiplication or division, but is unable to apply those concepts to fractions. Meanwhile, another student can compute with either whole numbers or fractions, but has difficulty with multiplication and division. If a mathematics test included four questions on whole numbers and four on fractions and each question required the student to employ a different type of computation function, both students would have the same aggregate score (50%). Yet, their identical scores would mask real differences in skills.

Standardized tests ignore the true complexity of human intelligence.

The assumption that test results can effectively be presented in the form of a unitary "score" presumes that the knowledge, skill or ability being measured is one-dimensional and that it tends to be distributed according to the statistical "normal" bell-shaped curve. Again research proves that this is not the case. Modern theories emphasize the complexity of human intelligence. Researchers have observed that knowledge, learning and thinking have multiple facets, and that a high level of development in one aspect does not necessarily indicate a high level of development in others.[16] Unitary test scores and linear scaling of scores ignore the true complexity and thus provide a deceptive picture of individual achievement, ability or skills.

For standardized achievement tests, problems also arise when test publishers use national test score averages ("norms") as reference points for interpreting student performance. These norms are developed by administering the test to a small group of students which, in theory, represents the national student population. Using these norms, schools can determine, for example, that a certain test score represents performance at the 65th national percentile, i.e. higher than 65% of all other students.

However, a recent study by Friends for Education raises substantial doubt about the validity of such national norms. Friends of Education "discovered that no state is below average at elementary level on any of the six major nationally normed, commercially available tests." They concluded that "these standardized, nationally normed achievement tests give children, parents, school systems, legislatures, and the press misleading reports on achievement levels. These tests allow all the states to claim to be above the national average."[17]

"Standardized, nationally normed achievement tests give children, parents, school systems, legislatures, and the press misleading reports on achievement levels."
--Friends for Education

Test Validation

Test validation determines the accuracy of a test in measuring what it claims to measure and the confidence one can have in conclusions drawn from test scores. However, validity cannot be determined in a vacuum. The validity of any standardized test depends entirely upon its context. Tests and test results are not valid for all purposes, but only for specific applications.

In theory, test validity depends upon three criteria:

* Do test results accurately measure the skills and knowledge actually taught to the students? (This is particularly important in tests used for diagnostic purposes.)

* Do test results accurately correlate with future performance in the subject being measured? (This is particularly important in tests used for admissions or course placement.)

* Do test results accurately correlate with the underlying characteristic or trait being measured? (This is particularly important in tests used to control promotions, program placement or graduations.)[18]

The validity of any standardized test depends entirely upon its context.

In reality, validation procedures focus almost exclusively on the first criteria (also known as content validity) and ignore the other two critically important procedures. Test developers (both commercial institutions and governmental agencies) generally validate a test's use by asking "experts" to make a qualitative judgment about the relationship between individual test items and the knowledge or skills they seek to measure.[19] In addition, "current standardized achievement tests are sometimes in part validated — for want of better criteria — by comparing scores on the tests with teachers' grades. . ." This latter approach to validation is particularly problematic, since "standardized tests originated [and are still promoted] in part as a substitute for teachers' judgments, which were deemed to be too subjective."[20] Moreover, the "experts" used to validate the tests are also inescapably subjective in their judgments.

"The term 'objective test' is a misnomer."
–Banesh Hoffman

Given this approach to validation, standardized tests reflect just as much subjective human judgment as do teachers' grades or schools' decisions to promote or graduate students. Why then are standardized tests called "objective tests" by test developers? As Banesh Hoffmann notes in *The Tyranny of Testing*, "the term 'objective test' is a misnomer. The objectivity resides not in the test as a whole but merely in the fact that no subjective element enters the *process* of grading once the key is decided upon."[21]

Test Reliability

Given the overall inadequacy of the validation process and the questionable assumptions implicit in their construction, it is not surprising that test results are often inconsistent and unreliable. The Congressional Budget Office, for example, noted that "one indication of the limitations of standardized tests is the often marked disparities in the results they yield."[22] Similarly, the Far West Laboratory for Educational Research and Development stated that "the older problem of test reliability remains: standardized tests rarely achieve reliability coefficients above .8, leaving considerable room for error in the measurement of any particular child's performance."[23] A reliability coefficient of .8 means that 20% of the individuals taking a test would score above or below a statistically determined level of test score variation if they were retested. Thus, scores (and the decisions made based on these scores) radically over- or underestimate the knowledge, skills or abilities possessed by at least one in every five students tested.

For younger children, the chances of a correct placement decision being made on the basis of test results is only 50%--the same odds as flipping a coin.

For younger children (those through grade 2), reliability is even worse — often only .5, according to Sue Bredekamp, director of early childhood programs for the National Association for the Education of Young Children (NAEYC). This means that the chances of a correct placement decision being made on the basis of test results is only 50% — "the same odds as flipping a coin."[24]

Test Administration

Educators, researchers and members of the public generally assume that standardized tests are administered in a standardized context under relatively uniform conditions. Anne Anastasi in *Psychological Testing* emphasized the importance of such a controlled setting: "Even apparently minor aspects of the testing situation may appreciably alter performance. . . In general, children are more susceptible to examiner and situational influences than are adults; in the examination of preschool children, the role of the examiner is especially crucial."[25]

In fact, recent research has demonstrated that tests are administered in far from "standard" conditions. One study concluded that "the actual context [of test administration] often includes confusion, anxiety, behavioral resistance, negative attitudes toward testing on the part of staff and students, lack of properly trained test examiners, developmentally or educationally immature children and other institutional problems that are endemic to many schools."[26] Given the reality of improper test administration, the use of these test results is likely to lead to educationally unsound decisions.

Recent research has demonstrated that tests are administered in far from "standard" conditions.

Moreover, many of the ideal "standard" conditions called for by test developers may actually place certain groups of students at a disadvantage. For example, the use of unfamiliar test examiners reduces test scores of low SES (socioeconomic status) and Black students, but does not affect the scores of high SES students. This factor alone can account for half of the difference in I.Q. test scores between low and high SES students.[27] Similarly, the time limitations associated with most standardized tests harm minorities. Generally, these groups appear to cope less effectively with the pressures inherent in a time-limited test than do White males.[28]

The use of unfamiliar test examiners reduces test scores of low SES and Black students, but does not affect the scores of high SES students.

III. THE IMPACT OF THE MISUSE OF TESTS BY THE PUBLIC SCHOOLS

"The adverse impact of an educational system based on tests may be far worse than anticipated. A lesson is to be learned from an English experiment [in 1863] which utilized tests as a basis for accountability. . . What transpired was nothing short of disaster. . . Almost all courses that were not addressed in the test were dropped and reading materials were limited to those that appeared on the tests. . . Some teachers quickly left the system and prospective teachers were reluctant to enter a profession that operated in this manner. Needless to say 'payment by results' was abandoned in fairly short order."[29]

- Mary Dilworth

Traditionally, standardized tests have been one of several educational tools used to assess student achievement and to diagnose their academic strengths and weaknesses. In recent years, American public schools, like their British counterparts more than a century ago, have begun to treat standardized tests as the all-purpose answer for promoting school improvement and ensuring school accountability. In the process, standardized tests have become the primary or sole criterion used by public schools for making a number of decisions affecting students, teachers and schools. In many schools, standardized tests serve as gatekeepers for:

American public schools have begun to treat standardized tests as the all-purpose answer for promoting school improvement and ensuring school accountability.

* assignment to special education or remedial programs;
* admission to gifted and talented or accelerated programs;
* grade promotions;
* high school graduation;
* merit pay awards to teachers;
* teacher certification and recertification;
* allocation of funds to schools or school systems; and
* school system certification and decertification.

Reliance on standardized tests as educational gatekeepers is growing. In just the last two years, the number of states using tests to determine student promotions increased from 8 to 12. Similarly, the number of states using tests to determine eligibility for high school graduation increased from 15 to 24.

The use of standardized tests as the primary or sole criteria for making any "high stakes" decision is reckless.

Given the limited range of skills and knowledge measured by standardized tests, the impact of race, ethnicity, income and gender on test results, and questions regarding their proper construction, validation and administration, the use of standardized tests as the primary or sole criteria for making any "high stakes" decision is reckless. Moreover, as standardized tests have become the all-powerful gatekeepers of American education, they have affected educational goals and curriculum, student progress and achievement and local control — and created a new set of problems in each area.

Impact on Educational Goals and Curriculum

Children go to school not just to learn basic academic skills, but also to develop the personal, intellectual and social skills to become happy, productive members of a democratic society. Unfortunately, the current emphasis on standardized tests threatens to undermine this educational diversity by forcing schools and teachers to focus on quantifiable skills at the expense of less easily quantifiable academic and non-academic abilities.

"Social, emotional, moral, and physical development and learning are virtually ignored in schools with mandated testing programs."
--NAEYC

This is particularly true for young children. As the National Association for the Education of Young Children (NAEYC) recently noted: "Many of the important skills that children need to acquire in early childhood — self-esteem, social competence, desire to learn, self-discipline — are not easily measured by standardized tests. As a result, social, emotional, moral, and physical development and learning are virtually ignored or given minor importance in schools with mandated testing programs."[30]

Many schools have embarked on a single-minded quest for higher test scores even though this severely narrows their curriculum:[31]

* Deborah Meier, a successful principal of a public school in Manhattan, noted that when synonyms and antonyms were dropped from the New York City test for word meaning, teachers promptly dropped academic material that stressed them. She also noted that students read "dozens of little paragraphs about which they then answer multiple choice questions" — an approach that duplicates the form of the standardized tests the students take in the spring.[32]

* Gerald Bracey, former Director of Research, Evaluation and Testing in the Virginia Department of Education, noted that some teachers did not teach their students how to add and subtract fractions because the state's minimum competency test included questions on multiplication and division of fractions, but not on their addition and subtraction.[33]

* In one Georgia school, the goal in essay writing is to produce "a five-paragraph argumentative essay written under a time limit on a topic about which the author may or may not have knowledge, ideas, or personal opinions". Not surprisingly, this exercise exactly matched the requirement for the Georgia Regents' Test essay exam.[34]

Sometimes, the curriculum is narrowed simply because "testing takes time, and preparing students for testing takes even more time. And all this time is time taken away from real teaching."[35] Unfortunately, a closer link between tests and curriculum has become a very conscious goal for some school systems.

"The curriculum falls in line with the test, and the test becomes the curriculum."
--P.S. Hlebowitsch

School systems in at least 13 states and the District of Columbia are seeking to "align" their curriculum so that students do not spend hours studying materials upon which they will never be tested regardless of the value or benefits which could be derived from that effort.[36] As a result, curriculum alignment "subordinates the process of curriculum development to external testing priorities, namely the state minimum-competency exam. Thus, the curriculum falls in line with the test, and, for all intents and purposes, the test becomes the curriculum."[37]

The educational price paid for allowing tests to dictate the curriculum can be a high one. Julia R. Palmer, Executive Director of the American Reading Council, recently wrote, "[T]he major barrier to teaching reading in a common-sense and pleasurable way is the nationally normed standardized second grade reading test." Ms. Palmer explained that the test questions force teachers and students to focus on "reading readiness" exercises and workbooks in their early grades and not on reading. As a result, many students become disenchanted with reading because they rarely get a chance to participate in it or to read anything of real interest to them.[38]

Just as curriculum has been narrowed, so too have textbooks. Diane Ravitch argues that "textbooks full of good literature began to disappear from American classrooms in the 1920's, when standardized tests were introduced. Appreciation of good literature gave way to emphasis on the 'mechanics' of reading."[39] Similarly, a recent report by the Council for Basic Education concluded that the emphasis on standardized tests and curriculum alignment are among the main causes of the increasingly poor quality of textbooks. The report noted that "instead of designing a book from the standpoint of its subject or its capacity to capture the children's imagination, editors are increasingly organizing elementary reading series around the content and time of standardized tests... As a result, much of what is in the textbooks is incomprehensible."[40]

The emphasis on standardized tests and curriculum alignment are among the main causes of the increasingly poor quality of textbooks.

By narrowing the curriculum, standardized tests are undermining many of the most important goals of the current school improvement movement.

Finally, by narrowing the curriculum, standardized tests are undermining many of the most important goals of the current school improvement movement. Current school reform efforts have sought to promote "higher-order thinking skills", imagination and creativity in American students. Yet standardized tests focus on basic skills, not critical thinking, reasoning or problem- solving. They emphasize the quick recognition of isolated facts, not the more profound integration of information and generation of ideas.[41] As Linda Darling-Hammond of the Rand Corporation concluded, "It's testing for the TV generation — superficial and passive. We don't ask if students can synthesize information, solve problems or think independently. We measure what they can recognize."[42]

Impact on Student Progress and Achievement

By controlling or compelling student placement in various educational programs, standardized tests perpetuate and even exacerbate existing inequities in educational services, particularly for minority and low-income students. Thus, standardized test results lead to larger numbers of racial and ethnic minorities being placed in special education and remedial education programs. For example, in 1984, 40% of Black five-year-olds in South Carolina failed standardized tests used to determine eligibility for kindergarten. These students were tracked into remedial classes on the basis of unreliable exams even before their formal schooling had begun.[43]

Standardized tests perpetuate and even exacerbate existing inequities in educational services, particularly for minority and low-income students.

Standardized tests also perpetuate the domination of white upper-middle class students in "advanced" classes. In New York City for example, IQ tests are used in some districts to place children in "gifted and talented" programs, creating White, upper-middle class enclaves in districts whose enrollment is dominated by racial and ethnic minorities.[44] Similarly, test results assign boys to advanced math and science programs and keep girls out. In the end, they both narrow the educational opportunities available to many segments of our student population and maintain the isolation of different racial and social groups and classes.

At the same time, standardized tests, particularly when used as promotional gates, can act as a powerful exclusionary device — again aimed disproportionately at minority and low-income students. Academic research has demonstrated that, for a student who has repeated a grade, the probability of dropping out prior to graduation increases by 40%.[45] Thus students who are not promoted because they fail an often unreliable and biased standardized test are considerably *more* likely to become high school dropouts.

Nor does the use of standardized tests affect only low-achieving students. High-achieving students are likely to be frustrated by a narrowed curriculum, which has been "dumbed down" in response to standardized tests, particularly minimum competency tests. These students too are likely to drop out in higher numbers.[46]

The most insidious effect of the overuse of standardized tests is on teachers' perceptions of their students. The existence of a "Pygmalion effect" as it relates to test results has long been a source of controversy. However, a 1984 study by Stephen Raudenbush has carefully documented its existence for students entering a new school (in this case, 7th grade students entering junior high school).[47] Where teachers have little information on students, conclusions about student knowledge, skills, and abilities based on often inaccurate and unreliable test results can become self- fulfilling prophesies.

Conclusions about student knowledge, skills, and abilities based on often inaccurate and unreliable test results can become self- fulfilling prophesies.

Impact on Local Control

Because standardized tests increasingly determine what is taught in the classroom, parents and other citizens are losing their traditional control over the public schools. This shift of power from local communities to state and national government reduces the level of input and influence available to both parents and teachers in the management of the schools. This, in turn, reduces "the responsiveness of schools to their clientele and so reduces the quality of education" available in those schools.[48]

Local control over the schools is also being lost to private organizations, namely the test developers. States and school districts have neither the expertise nor the resources to independently develop and validate the standardized tests that they need. Instead they turn to private testing companies, who design and market a tremendous variety of products. Even here, states and school systems have neither the skills nor the funds to adequately investigate claims by test developers regarding test validation or to review the test validation process.[49]

Local control over the schools is being lost to private organizations, namely the test developers.

Despite the significant and growing role their products play in educational decisions, testing manufacturers face little government regulation or supervision. Unlike other businesses, such as communications, food & drugs, transportation, and securities, there are virtually no regulatory structures at either the Federal or State level governing the billion dollar a year testing industry.

There are virtually no regulatory structures at either the Federal or State level governing the billion dollar a year testing industry.

States and school systems have neither the expertise nor the resources to adequately evaluate the claims made by the testing indus-

"It is practically impossible for a competent test technician or test consumer to make a thorough appraisal...because of the limited amount of trustworthy information supplied by the test publishers."

—Oscar K. Buros

try. Nor are they able to independently determine the reliability and validity of test results.[50] Even if the expertise and resources did exist, the secrecy which is rampant in the testing industry would likely prevent any effective outside evaluation. As the late Dr. Oscar K. Buros (editor of the *Mental Measurement Yearbook*) lamented, "It is practically impossible for a competent test technician or test consumer to make a thorough appraisal of the construction, validation, and use of standardized tests. . . because of the limited amount of trustworthy information supplied by the test publishers."[51]

IV. AN AGENDA FOR TEST REFORM

FairTest concurs with the National Academy of Education that "information on student progress, wisely interpreted, is of obvious value to the public, to educators and to policy-makers at all levels of government."[52] When properly constructed, validated, administered and used, standardized tests can serve as a tool in this effort.

Unfortunately, it has become all too obvious that standardized tests are *not* properly constructed, validated and administered. Moreover, their misuse is creating problems for students, teachers and the schools themselves. The question arises then: What should be done to reform tests and test use in the public schools?

In response to the misuse of standardized tests in U.S. society, FairTest has developed an agenda to answer this question. Our Testing Reform Agenda is guided by four basic principles:

Tests must be properly constructed, validated and administered.

• Tests should measure pertinent, not extraneous, knowledge differences between students. Questions must be relevant to the knowledge, abilities or skills being tested. Test items and instructions should be written clearly and accurately. The tests themselves should take into account the diversity of language, experience and perspective embodied in the test-taking population. If necessary, different tests should be used for different population groups to ensure the elimination of bias. At the same time, questions and scoring procedures should acknowledge the complexity and diversity of intelligence and individual development.

Test validation should ensure that the content of the test matches the content of what is taught, but test developers cannot stop at content validation. They must document assumptions about the relationship between test results and future performance. At the same time, they must demonstrate that test results are accurately related to the underlying knowledge, skills and abilities the test claims to measure.

Finally, those who develop and use standardized tests must ensure that the testing environment is both consistent for and supportive of all test-takers. Where the environment cannot be made standard and supportive, the only alternative is to refrain from testing. Moreover, the standard environment must not be constructed in a manner that creates disadvantages for particular students through artificial distractions or pressures.

Tests should be open.

• Public schools, test-takers and independent researchers should have access to the descriptive and statistical data needed to verify test publishers' claims regarding test construction and validation. This should include the release of questions used on previous tests, as well as data on test results (identified by race/ethnicity, gender, SES, geographical residence, and other demographic distinctions).

Test publishers have long argued against the release of old test questions. They have claimed that any large-scale release would require the development of a massive number of new items, thus increasing test development cost. This may not be the case. A recent study found that the release of old test questions did not affect test scores.[53] Thus the release of old test questions may not require such a large scale development of new questions.

Test users or independent public agencies should also fully investigate the claims of test publishers regarding the construction and validity of the tests. At the same time, they should disclose and monitor their own processes for test administration.

Tests should be viewed in the proper perspective.

• Both test developers and test users should work to ensure that test results are properly interpreted and employed by educators, policymakers, test-takers and the general public. As the 1974 *Standards for Educational and Psychological Tests* states: "A test score should be interpreted as an estimate of performance under a given set of circumstances. It should not be interpreted as some absolute characteristic of the examinee or as something permanent and generalizable to all other circumstances." This standard has too often been ignored by those who use test results. At a minimum, test scores should not be used as the sole or primary factor in "high stakes" educational decisions.

At the same time, test developers and test users must recognize that standardized tests are only limited measures of educational reality. Used alone, they present distorted pictures of what they seek to measure and often undermine the quality of education offered in our public schools. Both test developers and test users have the affirmative obligation to promote a proper, reasonable and limited use of standardized tests as one of a series of assessment mechanisms.

• Standardized multiple-choice tests can only measure a very limited range of student knowledge, abilities and skills. Emerging technologies provide new opportunities to expand our capability to more fully and accurately measure a greater range of knowledge, abilities and skills, while complementing the quantitative nature of current tests. These can be used to diagnose the strengths and weaknesses of students in order to help them learn, rather than to sort, stratify or segregate them. Despite claims that testing is now more advanced and scientific, Oscar Buros noted that "little progress has been made in the past fifty years — in fact in some areas, we are not doing as well. Except for the tremendous advances in electronic scoring, analysis and reporting of test results, we don't have a great deal to show for fifty years' work. Essentially, achievement tests are being constructed today in the same manner they were fifty years ago. . ."[54]

New assessment instruments must be developed as alternatives and supplements to standardized multiple-choice tests.

Although FairTest believes that those institutions that develop and use standardized tests have a primary obligation to reform tests and test use, government also has a major role to play. By establishing guidelines for the testing industry, requiring information on standardized tests to be made public, and analyzing test results to guard against bias, the government can go a long way toward improving the quality of tests and test use. More importantly, public agencies can set the standard for intelligent and proper use of test results. Too often, government is the biggest misuser of standardized test results.

Unfortunately too many policymakers and educators have ignored the complexities of testing issues and the obvious limitations they place upon standardized test use. Instead, they have been seduced by the promise of simplicity and objectivity. The price which has been paid by our schools and our children for this infatuation with tests is high. Unless Americans act now to limit and reform the use of standardized tests in the schools, that price will continue to increase.

NOTES

[1] W. Haney & G. Madaus. "Effects of Standardized Testing and the Future of the National Assessment of Educational Progress", *Working Paper for the NAEP Study Group* (1986), p. 5.

[2] National Academy of Education. *The Nation's Report Card: Improving the Assessment of Student Achievement*, Review of the Alexander/James Study Group Report. (National Academy of Education, 1987), p. 47.

[3] A.K. Wigdor & W.R. Garner (eds.). *Ability Testing: Uses, Consequences & Controversies.* (National Academy Press, 1982), pp. 252-253.

[4] Education Research Service, Inc. *Kindergarten Programs & Practices in Public Schools.* (1986).

[5] C. Pipho. "Tracking the Reforms: Part 5 - Testing", *Education Week* (May 22, 1985), p. 19.

[6] Congressional Budget Office. *Trends in Educational Achievement.* (Government Printing Office: Washington, D.C.: April 1986), p. 20.

[7] J. Gannon. "Future of Affirmative Action in Law School Admissions". (Unpublished paper, 1980).

[8] M.R. Hoover, R.L. Politzer & O. Taylor. "Bias in Reading Tests for Black Language Speakers: A Sociolinguistic Perspective", *Negro Educational Review* (April-July 1987), pp. 81 - 98.

[9] J. Loewen. "Possible Causes of Lower Black Scores on Aptitude Tests". (Unpublished research report, 1980).

[10] O.L. Taylor & D.L. Lee. "Standardized Tests and African-American Children: Communication and Language Issues", *Negro Educational Review* (April-July 1987), pp. 67 - 80.

[11] "The 'Culture Fair' WISC-R I.Q. Test", *The Testing Digest* (Spring 1981), p. 21. See also issues of *FairTest Examiner.*

[12] For research on Hispanics, see A.P. Schmitt. "Unexpected Differential Item Performance of Hispanic Examinees on the SAT-Verbal, Forms 3FSAO8 and 3GSAO8". (Unpublished statistical report of the Educational Testing Service, 1986). Dr. Schmitt, an ETS researcher, concluded that Mexican-American students scored significantly higher than expected on a reading comprehension passage concerned with lifestyle changes in Mexican-American families. For research on Blacks, see Hoover, Politzer & Taylor (1987), p. 83. Dr. Darlene Williams found that "the use of pictures showing Blacks and related to Black culture raised IQ scores for all Black children." For research on females, see J.W. Loewen, P. Rosser & J. Katzman. "Gender Bias in SAT Items." Paper presented at the AERA Annual Convention (New Orleans, La., April 5, 1988).

[13] The mathematics section of the WISC-R test, for example, includes 8 questions about 13 boys or men who save money on purchases, trade fairly, cleverly divide their efforts and money and work at jobs, compared to only one question featuring a girl who loses her hair ribbon.

[14] Taylor & Lee. (1987).

[15] D. Meier. "Why Reading Tests Don't Teach Reading", *Dissent* (Winter 1982-83). Meier notes that questions correctly answered by "low-performing" students, but not by "high-performing" students are removed from test use as inaccurate measures of ability or achievement. This ignores the possibility that these questions reflect differing language or cognitive styles possessed by "low-performing" students.

[16] H. Gardner. *Frames of Mind: The Theory of Multiple Intelligences.* (Basic Books, 1985).

[17] J. J. Cannell. *Nationally Normed Elementary Achievement Testing in America's Public Schools.* (Friends for Education: Daniels, West Virginia, 1987), p. 6.

[18] G. Madaus & D. Pullin. "Questions to Ask When Evaluating a High-Stakes Testing Program", *NCAS Backgrounder* (June 1987).

[19] Madaus & Pullin. (1987).

[20] Congressional Budget Office. *Educational Achievement: Explanations and Implications of Recent Trends* (Government Printing Office: Washington, D.C., August 1987).

[21] B. Hoffmann. *The Tyranny of Testing.* (Crowell-Collier, 1962), pp. 60 - 61.

[22] Congressional Budget Office. (1986), p. 10. See also the 1987 report by CBO, pp. 10 - 11, which concluded that "one of the most serious mistakes made by some analysts attempting to explain recent achievement trends . . . has been to assume that patterns evident in the scores of one test will appear in others as well."

[23] D.E. Mitchell & D. Zalles. "Student Assessment Programs in the West", *Policy Briefs* # 4 (Far West Laboratory for Educational Research & Development, Spring 1987), p. 2.

[24] "Mass Academic Testing of Young Children Should Stop, Groups Argue," *Education Week* (March 25, 1988), p. 5.

[25] A. Anastasi. *Psychological Testing.* (Macmillan: New York City, N.Y., 1982).

[26] K. Wodtke, F. Harper, M. Schommer & P. Brunelli. "Social Context Effects in Early School Testing: An Observational Study of the Testing Process". (1985), p. 28.

[27] D. Fuchs & L.S. Fuchs. "Test Procedure Bias: A Meta-Analysis of Examiner Familiarity Effects", *Review of Educational Research* (Summer 1986), pp. 243-262. See also "Test Conditions Can Harm Minority-Group Children", *The Chronicle of Higher Education* (November 18, 1987), p. A15.

[28] G.I. Maeroff. "Reading Test Time Limits Are Criticized," *New York Times* (January 29, 1985) citing a study an Wingate High School in Brooklyn, N.Y. which concluded that Black students significantly improved their test scores when the test was administered without a time limit.

[29] M.E. Dilworth. *Teachers' Totter: A Report on Teacher Certification Issues.* (Institute for the Study of Educational Policy, 1984), pp. 33-34.

[30] National Association for the Education of Young Children. "Testing of Young

Children: Concerns and Cautions" (pamphlet), (1988). See also, "NAEYC Position Statement on Standardized Testing of Young Children 3 Through 8 Years of Age," *Young Children* (1988), pp. 42 - 47.

[31] G. F. Madaus. *Testing and the Curriculum: From Compliant Servant to Dictatorial Master.* (Center for the Study of Testing, Evaluation and Educational Policy: Chestnut Hill, Ma., 1986). See also, H.C. Rudman. "Testing Beyond Minimums," *ASAP Notes* (Occasional Paper # 5, 1985), p. 1 - 36.

[32] G.F. Madaus. "Test Scores as Administrative Mechanisms in Educational Policy", *Phi Delta Kappan* (May 1985), p. 616.

[33] "Some 'Teach' to the Test", *The [Newport News, VA] Daily Press* (June 15, 1987), p. C1.

[34] Haney & Madaus. (1986), p. 19.

[35] A.E. Wise. "Legislated Learning Revisited", *Phi Delta Kappan* (January 1988), pp. 330. See also, D.W. Dorr-Bremme & J.L. Herman. *Assessing Student Achievement: A Profile of Classroom Practices.* (Center for the Study of Evaluation, UCLA: Los Angeles, Ca., 1986) which concluded, based upon a nationally representative sample of 114 school districts, that "preparations for a test can begin days or even weeks before the test is given". They cited one district where students spend only 55% of the time allotted to tests actually taking them. The remainder of the time is spent on test review, practice and preparation. They also observed that "for each hour that students spend taking tests, teachers seem to spend two to three more".

[36] L. Olson. "Districts Turn to Nonprofit Group for Help in 'Realigning' Curricula to Parallel Tests", *Education Week* (October 18, 1987), p. 1 & 19. The Education Products Information Exchange (EPIE) has worked on "curriculum alignment" with school systems in thirteen states including Alaska, California, Connecticut, the District of Columbia, Georgia, Kansas, Missouri, New Jersey, New York, North Carolina, Ohio, Pennsylvania, Rhode Island, and Tennessee.

[37] P.S. Hlebowitsh. Letter to the editor, *Education Week* (November 18, 1987), p. 21.

[38] J.R. Palmer. Letter to the editor, *New York Times* (December 14, 1987). See also, J.T. Guthrie. *Indicators of Reading Education.* (Center for Policy Research in Education: New Brunswick, N.J., 1988), which concludes that the strengthening of students' reading skills goes hand-in-hand with finding better ways to measure reading achievement. The primary shortcoming of reading tests is that they don't reflect the complexity of the reading process.

[39] E.B. Fiske. "America's Test Mania", *New York Times* (April 10, 1988), Section 12, p. 20.

[40] R. Rothman. "Textbook Rules Have Backfired, Report Contends", *Education Week* (April 20, 1988), p. 1.

[41] A. Bastian, et.al. *Choosing Equality: The Case for Democratic Schooling.* (Temple University Press: Philadelphia, PA., 1986), p. 73.

[42] Fiske. *New York Times* (April 10, 1988), p. 20.

[43] South Carolina Department of Education. *Report on the Implementation of the Basic Skills Assessment Program, 1984- 1985*. (December 1985), p. 8.

[44] A. Cook, Community Studies, Inc., New York City, NY. (Personal communication, April 1988).

[45] Massachusetts Advocacy Center. Memorandum to the Boston School Committee (June 19, 1987) quoting from Office of Educational Assessment, New York City Board of Education, "Evaluation Update on the Effect of the Promotional Policy Program" (November 12, 1986). See also, M.L. Smith and L.A. Shepard. "What Doesn't Work: Explaining Policies of Retention in the Early Grades", *Phi Delta Kappan* (October 1987), pp. 129 - 134.

[46] "Student Competency Exams Present Major Barrier to Minority Students", *Education Daily* (August 27, 1987), p. 3.

[47] S. Raudenbush. "Magnitude of Teacher Expectancy Effects on Pupil IQ As a Function of the Credibility of Expectancy Induction — A Synthesis of Findings From 18 Experiments", *Journal of Educational Psychology* (January 1984), pp. 85 - 97.

[48] A.E. Wise. (January 1988), pp. 328 - 333. See also, A. Porter. "Indicators: Objective Data or Political Tool?", *Phi Delta Kappan* (March 1988), pp. 503 - 508.

[49] Madaus & Pullin. (1987), p. 3 - 4.

[50] Madaus & Pullin. (1987), p. 3.

[51] O. Buros. "Fifty Years in Testing: Some Reminiscenses, Criticisms, and Suggestions", *Educational Researcher* (Summer 1978), p. 14.

[52] National Academy of Education. (1987), p. 47.

[53] W. Haney. "Test Reasoning and Reasoning about Testing", *Review of Educational Research* (Winter 1984), p. 628.

[54] W. Haney. (1984), p. 635. Quoting O.K. Buros.

APPENDIX

DESCRIPTION OF FAIRTEST SURVEY AND RESULTS

Study Methodology

In mid-1987, FairTest staff conducted a series of telephone interviews with officials from all 50 state departments of education, from the District of Columbia school district and from 56 sample school districts in 38 states. Interviews of state officials focused on standardized tests administered by the public schools to fulfill testing mandates established by the state, while interviews of school district officials focused on tests administered to fulfill testing mandates established by the district. All interviews sought responses to three basic questions:

* How many tests were administered by the public schools to fulfill the state or local testing mandates?

* Which standardized tests were used to fulfill the state or local testing mandates?

* For what purposes did the state or school district mandate standardized tests?

The responses collected through these interviews related entirely to the use of standardized achievement, competency or basic skills tests. Public schools also use many other standardized exams, including IQ tests, behavioral tests, readiness tests for young children, and placement tests for special education, remedial education and bilingual education programs. However, the use of these tests varies considerably among schools, even within the same districts, and records of their use appear to be unreliable or nonexistent. Thus the study results reflect only a portion of the standardized tests actually administered to students by the public schools.

Study Results

State-Level Testing Mandates. During the 1986-87 school year, schools in 42 states and the District of Columbia administered over 17.5 million standardized achievement, competency and basic skills tests to almost 36.3 million students to fulfill state testing mandates — a rate of almost one test for every two students. This rate varied considerably among the states however. A detailed listing of the number of standardized tests administered by the public schools in each state to fulfill state testing mandates is presented in Table 1.

On average, schools in the South administered standardized tests to fulfill state mandates at much higher rates than schools in the remainder of the nation. Schools in the 11 Southern states (Alabama, Arkansas, Florida, Georgia, Kentucky, Louisiana, Mississippi, North Carolina, South Carolina, Tennessee and Virginia) administered more than 6.7 million tests to over 9.3 million students — a rate of one test for every 1.3 students. Schools in the remaining 31 states administered tests at about half that rate — one test for every 2.5 students. In fact, schools in all but one Southern state (Florida) administered tests at a rate higher than the average for the remainder of the Nation. Moreover, the five states with the highest rates of test administration were all in the South. Virginia, Kentucky, North Carolina, Alabama, and Georgia all administered about one test per student.

Outside the South, no clear patterns of test use to fulfill state mandates emerged. Although schools in New Mexico, a state with a minority population in the public schools above the national average, reported a high rate of standardized test use, so did schools in Utah, a state with a relatively low minority population. Conversely, schools in states like Texas and California, with relatively high minority populations in the public schools, reported rates almost equal to Wisconsin and Kansas, states with relatively low minority populations.

A clear pattern did emerge among the eight states (Alaska, Iowa, Minnesota, Montana, North Dakota, Ohio, Vermont and Wyoming) which did not have any state testing mandates. Seven of these eight states have minority student populations significantly below the national average. While minorities make up more than one-quarter of the public school students nationwide, the minority student populations in these seven states range from a high of 15% in Ohio to a low of 1% in Vermont. Most in fact have minority populations that are less than 10% of the total student population. The one exception is Alaska, with a large Native American population, and an overall minority population just above the national average. However, Alaska will also implement a state testing mandate beginning in the 1987-88 school year.

District-Level Testing

During the 1986-87 school year, schools in the 56 sample school districts administered more than 7.8 million standardized tests to over 5.7 million students to meet local testing mandates — a rate of one and one-third tests for every student. However, the rate of test administration among districts varies even more than it did among

different states. A detailed listing of the number of standardized tests administered by the public schools in each of the sample school districts to fulfill local testing mandates is presented in Table 2.

The schools in Newark, New Jersey are excluded from the discussion in the following paragraph due to their extremely high reliance on standardized tests. During the 1986-87 school year, schools in Newark administered over 600,000 standardized tests to about 67,000 students — a rate of more than 9 tests per student. This rate was more than three times that of the next highest school system.

Excluding the Newark school system, the rate of test administration ranges from a high of almost 3 tests per student in Cleveland to a low of only one test per 12.5 students in Fairfax County, Virginia. Overall, six districts reported administering more than 2 tests per students per year, while seven others reported administering less than 1 test for every 2 students.

To analyze the variations among the different rates of tests administered by the different districts, districts were categorized by size. Three categories were created: large districts (with student populations exceeding 100,000); medium- sized districts (with populatio s between 25,000 and 100,000); and small districts (with populations below 25,000). The results from Newark were excluded from these calculations. On average, large districts administered standardized tests to fulfill local testing mandates at a rate (1.38 tests per student) 25% higher than that of medium-sized districts (1.11 tests per student). The rate of me .um-sized districts, in turn, exceeded that of small districts (0.88 tests per student) by almost the same proportion.

The estimate that schools administered almost 38.9 million standardized achievement, competency and basic skills tests to 39.8 million students during the 1986-87 school year to fulfill local testing mandates is based upon these average variations and the distribution of students among school districts of different sizes. (See Table 3 for the detailed computations.) Combining this figure with the tests administered to fulfill state testing mandates produces an estimate of 56.4 million standardized achievement-type tests administered by the public schools during the 1986-87 school year, or 1.4 tests for every public school student.

Additional Surveys on Test Use

Although the FairTest survey focused on the use of standardized achievement, competency and basic skills tests in the public schools, public schools also use standardized tests for a variety of other purposes. These include:

* tests administered to gifted, disadvantaged, handicapped and limited-English proficient student for placement into or graduation from gifted & talented, compensatory education, special education and bilingual education programs;
* college admissions tests administered to high school juniors and seniors;

* admissions tests administered to students seeking to enroll in particular secondary schools in certain communities;

* screening tests administered to kindergarten and pre- kindergarten students;

* GED (General Education Development) tests administered to individuals who did not complete high school; and

* tests administered to randomly selected samples of students as part of the U.S. Department of Education's National Assessment of Educational Progress.

In addition, some school systems continue to administer IQ tests to their entire student populations, although most school districts administer IQ tests only for placement purposes.

As noted previously, information on the number of other standardized tests administered was neither as specific nor as reliable as the information gathered for achievement, competency and basic skills tests. As a result, it is not possible to compute specific totals on the use of these tests. From a variety of sources, however, general figures can be obtained:

* Test publishers have reported that the total number of college and secondary school admissions, GED and NAEP tests administered to elementary and secondary school students was between 6 and 7 million.

* A survey conducted in the early 1980's indicated that students in compensatory education and special education programs were tested two to three times as often as their peers in mainstream

education program.[1] Since the mainstream student is administered 1.4 standardized tests per year (according to our survey), compensatory and special education students are administered between 3 and 4 standardized tests per year. Given that over 10 million students participate in Federally- funded compensatory and special education programs, an estimated 30 to 40 million additional standardized tests are administered to these students.

* A survey conducted in 1985 concluded that almost half of the kindergarten and pre-kindergarten students in the public schools were administered screening tests.[2] Based upon the 1985 kindergarten enrollment, this means that between 1.5 and 1.75 million tests were administered to these children.

This total (which still excludes tests administered to gifted and limited-English proficient students and some proportion of I.Q. testing) yields an additional 37.5 to 48.75 million tests.

Summary

The results of the FairTest survey revealed that at least 56.4 million standardized achievement-type tests were administered to public school students in the 1986 - 87 school year. An additional 37.5 to 48.75 million standardized tests were administered by the public schools to their students for other purposes during that year. Based on these two estimates, between 94 million and 105 million standardized tests were administered to 39.8 million students in the public schools.

NOTES

[1] A.K. Wigdor & W.R. Garner (eds.). *Ability Testing: Uses, Consequences and Controversies* (National Academy Press, 1982) pp. 252-253.

[2] Educational Research Services, Inc. Kindergarten Programs & Practices in the Public Schools (1986).

TABLE 1. Number of Standardized Tests Administered in the Public Schools To Fulfill State Mandates, By State (1986-1987 School Year).

STATE	NUMBER OF TESTS	SCHOOL ENROLLMENT	STATE-MANDATED TESTS PER STUDENT
Alabama	720,000	733,735	0.98
Alaska	0	107,973	0.00
Arizona	480,000	534,538	0.90
Arkansas	179,000	437,438	0.41
California	1,420,000	4,377,989	0.32
Colorado	208,800	558,415	0.37
Connecticut	99,000	468,847	0.21
Delaware	60,000	94,410	0.64
District of Columbia	119,000	85,612	1.39
Florida	437,352	1,607,320	0.27
Georgia	1,020,000	1,096,425	0.93
Hawaii	75,000	164,640	0.46
Idaho	30,000	208,391	0.14
Illinois	418,000	1,825,185	0.23
Indiana	211,000	966,780	0.22
Iowa	0	481,286	0.00
Kansas	135,000	416,091	0.32
Kentucky	650,000	642,778	1.01
Louisiana	390,000	795,188	0.49
Maine	48,000	211,752	0.23
Maryland	168,000	675,747	0.25
Massachusetts	410,000	833,918	0.49
Michigan	326,285	1,681,880	0.19
Minnesota	0	711,134	0.00
Mississippi	240,000	498,639	0.48
Missouri	664,000	800,606	0.83
Montana	0	153,327	0.00
Nebraska	20,508	267,139	0.08
Nevada	55,000	161,239	0.34
New Hampshire	30,000	163,717	0.18
New Jersey	630,000	1,107,467	0.57
New Mexico	80,000	281,943	0.28
New York	1,691,000	2,607,719	0.65
North Carolina	1,085,000	1,085,248	1.00
North Dakota	0	118,703	0.00
Ohio	0	1,793,508	0.00
Oklahoma	234,000	593,183	0.39
Oregon	15,000	449,307	0.03
Pennsylvania	538,212	1,674,161	0.32
Rhode Island	40,000	134,126	0.30
South Carolina	450,000	611,629	0.74
South Dakota	50,000	125,458	0.40
Tennessee	500,000	818,073	0.61
Texas	1,500,000	3,209,515	0.47
Utah	258,000	415,994	0.62
Vermont	0	92,112	0.00
Virginia	1,069,000	975,135	1.10
Washington	182,000	761,428	0.24
West Virginia	200,000	351,837	0.57
Wisconsin	309,000	767,819	0.40
Wyoming	0	100,955	0.00
UNITED STATES	17,545,159	39,837,459	0.44

Table 2. Standardized Tests Administered in the Public Schools To Fulfill State Mandates (1986-1987 School Year)..

SCHOOL DISTRICT	NUMBER OF TESTS	SCHOOL ENROLLMENT	DISTRICT-MANDATED TESTS PER STUDENT
CALIFORNIA			
Los Angeles	949,899	540,903	1.76
San Juan Unified	33,000	44,186	0.75
San Francisco	21,000	58,378	0.36
San Diego City Unified	34,040	110,631	0.31
COLORADO			
Jefferson County	52,500	76,000	0.69
CONNECTICUT			
Greenwich	5,720	6,772	0.84
FLORIDA			
Hillsborough County (Tampa)	98,000	115,323	0.85
Pinellas County (Clearwater)	112,000	85,339	1.31
Duvall County (Jacksonville)	190,000	99,512	1.91
Dade County (Miami)	230,000	250,000	0.92
Broward Cty. (Ft. Lauderdale)	336,000	129,478	2.60
GEORGIA			
Fulton County (Atlanta)	13,000	35,523	0.37
DeKalb County (Decatur)	15,000	66,000	0.23
ILLINOIS			
Chicago	468,544	435,000	1.08
INDIANA			
Indianapolis	66,600	50,600	1.32
IOWA			
Des Moines Independent	44,500	30,000	1.48
KANSAS			
Wichita	65,729	44,729	1.47
KENTUCKY			
Jefferson County (Louisville)	47,000	95,020	0.49
LOUISIANA			
Orleans Parish (New Orleans)	84,000	84,000	1.00
MARYLAND			
Baltimore City	275,800	120,000	2.30
Prince Georges County	175,000	103,000	1.70
MASSACHUSETTS			
Boston	60,000	55,000	1.09
Brookline	3,500	5,400	0.65
MICHIGAN			
Detroit City	269,000	200,000	1.35
MINNESOTA			
Minneapolis	39,712	32,274	1.23

SCHOOL DISTRICT	NUMBER OF TESTS	SCHOOL ENROLLMENT	DISTRICT-MANDATED TESTS PER STUDENT
MISSOURI			
St. Louis	119,327	48,800	2.45
MONTANA			
Missoula	6,990	5,640	1.24
NEVADA			
Las Vegas	87,684	95,000	0.93
NEW HAMPSHIRE			
Concord	7,296	5,000	1.46
NEW JERSEY			
Newark	603,000	67,000	9.00
NEW MEXICO			
Albuquerque	133,320	80,000	1.67
NEW YORK			
New York City	1,133,000	924,123	1.23
Rochester	35,000	34,696	1.01
Buffalo City	63,000	44,707	1.41
NORTH CAROLINA			
Mecklenburg Cty. (Charlotte)	18,000	72,162	0.25
Wake County (Raleigh)	40,500	58,213	0.70
NORTH DAKOTA			
Fargo	6,000	9,200	0.65
OHIO			
Cincinnati	123,800	53,000	2.34
Akron	20,520	34,000	0.60
Cleveland	219,000	75,000	2.92
OKLAHOMA			
Oklahoma City	31,300	40,000	0.78
OREGON			
Portland	30,000	52,000	0.58
PENNSYLVANIA			
Philadelphia	400,000	200,000	2.00
SOUTH CAROLINA			
Greenville County	37,000	53,000	0.70
TENNESSEE			
Memphis City	188,200	106,000	1.76
TEXAS			
Dallas Independent	217,584	127,584	1.71
Houston Independent	250,000	193,702	1.29
UTAH			
Salt Lake City	56,000	72,000	0.78

SCHOOL DISTRICT	NUMBER OF TESTS	SCHOOL ENROLLMENT	DISTRICT-MANDATED TESTS PER STUDENT
VERMONT			
Burlington	2,212	3,800	0.58
VIRGINIA			
Fairfax County	10,000	124,631	0.08
Virginia Beach City	0	54,870	0
Prince William County	105,000	57,213	1.84
WASHINGTON			
Seattle	41,500	44,000	0.94
WEST VIRGINIA			
Kanawha County (Charleston)	20,300	37,399	0.54
WISCONSIN			
Milwaukee	64,500	96,387	0.67
WYOMING			
Laramie County (Cheyenne)	11,000	13,000	0.84

TABLE 3. COMPUTATION OF ESTIMATE OF STANDARDIZED TESTS ADMINISTERED BY PUBLIC SCHOOLS TO FULFILL STATE OR LOCAL TESTING MANDATES (1986-87 SCHOOL YEAR).

	SURVEYED			ESTIMATED		TOTAL	
TYPE OF SYSTEM	STUDENTS	TESTS	RATE	STUDENTS	TESTS	STUDENTS	TESTS
LARGE (OVER 100,000)	3,680,375	5,035,067	1.37	—	—	3,680,375	5,035,067
MEDIUM (25,000 TO 100,000)	2,026,008	2,790,792	1.11	4,674,000	5,188,000	6,700,008	7,978,792
SMALL (LESS THAN 25,000)	48,812	42,718	0.88	29,408,000	25,879,000	29,456,812	25,921,718

TESTS ADMINISTERED TO FULFILL -	TESTS IN SURVEY	ESTIMATED TESTS	TOTAL TESTS
STATE MANDATES	17,545,159	—	17,545,159
LOCAL MANDATES	7,868,577	31,067,000	38,935,577
ALL MANDATES	25,413,736	31,067,000	56,480,736

NOTE: The computation of the testing rate in medium-sized school systems excludes the unusually high testing rate of Newark, New Jersey (which is three times higher than the next highest system.) The estimated number of tests administered in medium and small school systems was computed using the actual rate administered for medium and small systems based on survey results.

ANNOTATED BIBLIOGRAPHY

1. *GENERAL DISCUSSIONS OF TESTING*

Bastian, A., et.al. *Choosing Equality: The Case for Democratic Schooling.* (Temple University Press: Philadelphia, PA, 1986).

This book concludes that democratizing schools so as to meet the needs of all students ought to be the focus of school reform efforts. Standardized tests are seen as one of the means used to prevent equality and quality for all.

FairTest. *FairTest Examiner.*

This newsletter surveys developments in testing and testing reform, including grade school, IQ and related tests.

(This is available from FairTest, Box 1272, Cambridge, MA 02238.)

FairTest. *Update on K - 12 Testing.*

A biannual selection of newspaper and journal articles on grade school, IQ and related tests and testing reform efforts. This document complements information released in the *Examiner* by providing more in-depth review. This is one of seven *Updates* provided by FairTest on various aspects of testing.

(This is available from FairTest, Box 1272, Cambridge, MA 02238.)

Fiske, E. "America's Test Mania". *New York Times* (April 10, 1988) Section 12, pp. 16 - 20.

This article contains a general overview of the growth of standardized testing in America's public schools and the resulting problems. It discusses the various uses being made of standardized tests, efforts by schools to improve test scores and criticisms and limitations of standardized tests. It also suggests that standardized tests will undermine the goals of the current school reform movement.

Haney, W. "Test Reasoning and Reasoning about Testing", *Review of Educational Research* (Winter 1984), p. 628.

This article provides a detailed history of the use of standardized tests in America (divided into three eras: pre-WWI; WWI to 1950;

and 1950 to present). In particular, it chronicles the recent growth in test use and test criticism. The article includes a brief discussion of the intensity of use of standardized tests in the public schools and a longer discussion of the types of uses. The article concludes by suggesting avenues for further research and development in testing.

Negro Educational Review (Special Issue). **"Testing African American Students," (April - July 1987).**

This issue contains numerous articles on testing, including discussions of: psychometric, language and cultural biases against Blacks (particularly working class Blacks) and other minorities; IQ testing and the *Larry P.* decision; and alternatives to standardized tests.

(This is available from Box 2895, Jacksonville, FL 32203.)

Phi Delta Kappan (Special Issue). **"What is the Proper Role of Testing," (May 1985), pp. 599 - 639.**

The nine feature articles focus on the use, problems and impact of standardized testing in the public schools. These articles cover: the misuse of SAT scores in assessing the quality of American education; the popularity of standardized tests; the impact of testing on pedagogy and instruction; problems with teacher testing; the impact of testing on educational equity; and alternatives to multiple-choice writing tests. One article also provides a description of the use of standardized tests to "drive the curriculum" in three states and one school district.

2. *USE OF STANDARDIZED TESTS IN THE PUBLIC SCHOOLS*

ipho, C. **"Tracking the Reforms: Part 5 - Testing", *Education Week* (May 22, 1985), p. 19.**

One in a series of articles discussing the state education reform efforts of 1983 - 1985. This article includes a detailed listing of state usage of standardized tests as part of the reforms.

3. *PROBLEMS WITH STANDARDIZED TESTS*

Congressional Budget Office. ***Trends in Educational Achievement* (April 1986).**

The first of two CBO reports on educational achievement. It includes a general discussion of the ways in which standardized test results are used to measure educational achievement and the problems with such uses. The report warns against overreliance on test results in evaluating trends in educational achievement.

(This document is available from the Superintendent of Documents, U.S. Government Printing Office, Washington, D.C, 20402.)

Haney, W. & G. Madaus.

"Effects of Standardized Testing and the Future of the National Assessment of Educational Progress", *Working Paper for the NAEP Study Group* (1986), p. 5.

This article, commissioned by the Alexander/James Study Group on NAEP, documents the increasing attention paid to testing compared with curriculum in educational literature. It then discusses four broad issues which determine the impact of testing: the subject of the test; how tests are referenced; the sources of the tests; and their rewards or sanctions. It next lists seven major problems regarding the impact of tests on individuals and on schooling. The authors suggest efforts that can minimize the negative impact and maximize the possible benefits of testing. Two sections specifically discuss NAEP and its future.

(This document is available from ERIC Document Reproductions as Document Number ED-279-680.)

Madaus, G. & D. Pullin.

"Questions to Ask When Evaluating a High- Stakes Testing Program", *NCAS Backgrounder* (June 1987).

This report, presented in a question-and-answer format focuses on the use of standardized tests in "high stakes" educational situations (i.e. where significant sanctions or rewards are associated with the test). The report lists several "high stakes" uses of tests. The article discusses the ways in which tests are designed or selected. The main section of the article examines the potential structure of testing programs, possible conclusions that will be drawn from their results and their likely impact on students and schools.

(This document is available from NCAS, 100 Boylston St., Boston, MA 02116.)

National Association for the Education of Young Children (NAEYC).

"Testing of Young Children: Concerns and Cautions" (pamphlet). (NAEYC: Washington, D.C., 1988).

This pamphlet discusses the potentially harmful impact on young children of standardized testing. The pamphlet also describes proper uses of standardized tests and suggest how school can "help ensure that all children get off to a sound start in kindergarten, first, and second grade."

(This document is available from NAEYC, 1834 Connecticut Avenue, N.W., Washington, D.C. 20009.)

A. *Test Bias*

Hoover, M.R., R.L. Politzer & O. Taylor.

"Bias in Reading Tests for Black Language Speakers: A Sociolinguistic Perspective". ***Negro Educational Review,*** **(April - July 1987), pp. 81 - 98.**

This article contains a detailed discussion of language-related bias in standardized tests against speakers of non-standard English, including phonological (sound), syntactical (structural), and lexical (work choice and vocabulary) biases. The consequences of bias include school program misplacement and tracking resulting in inadequate education for those who are not White middle- to upper-class. Eliminating these test biases is important for reducing educational and societal biases against working class and minority children.

(This article is available from N.E.R. Box 2895, Jacksonville, FL 32203).

Schmitt, A.P.

"Unexpected Differential Item Performance of Hispanic Examinees on the SAT-Verbal, Forms 3FSAO8 and 3GSAO8" (Unpublished statistical report of the Educational Testing Service, 1986).

This report analyzes student responses on specific SAT questions and concludes that Hispanic students scored significantly higher than expected on a reading comprehension passage related to subjects in which they are likely to have an interest (e.g. lifestyle changes in Hispanic families).

(This article is available from E.T.S., Rosedale Rd., Lawrence, N.J. 08541.)

Taylor, O. & D.L. Lee.

"Standardized Tests and African Americans: Communication and Language Issues". ***Negro Educational Review,*** **(April - July 1987), pp. 67 - 80.**

This article contains a detailed discussion of sources and kinds of cultural and language bias in standardized tests. These biases cause African-American (particularly working-class Blacks) and others who are not middle- to upper-class Whites to be invalidly assessed: "At times. . . the results fail to accurately represent actual abilities." In conclusion, ". . . the very assumptions and paradigms upon which most standardized tests are based need to be revised."

(This article is available from N.E.R., Box 2895, Jacksonville, FL 32203).

B. *Test Validation & Reliability*

Cannell, J.J. — *Nationally Normed Elementary Achievement Testing in America's Public Schools: How All Fifty States Are Above the National Average.* (Friends for Education: Daniels, W.Va., 1987).

This reports describes the results of a nationwide survey on achievement test scores. It notes that no state average score at the elementary level was below the national norm on any of the six most popular achievement tests. The report concludes that this results from improper norming of the tests and from teaching to the tests.

(This article is available from Friends for Education, Box 358, Daniels, W.Va. 25832.)

Congressional Budget Office. *Educational Achievement: Explanations and Implications of Recent Trends* (August 1987).

The second of two CBO reports on educational achievement. This report includes a general discussion of the problems and limitations of standardized tests and particularly the problems around validity and reliability of the tests.

C. *Test Administration*

Fuchs, D. & L.S. Fuchs. — "Test Procedure Bias: A Meta-Analysis of Examiner Familiarity Effects", *Review of Educational Research* (Summer 1986), pp. 243-262.

The authors of this report analyzed data from 22 controlled studies involving 1489 subjects. Their analysis discovered that familiarity with the test examiner had different impacts on test- takers. For low-income and Black students, the use of unfamiliar examiners reduces test scores, but does not affect the scores of upper-income students.

Wodtke, K., F. Harper, M. Schommer & P. Brunelli. — "Social Context Effects in Early School Testing: An Observational Study of the Testing Process" (1985).

Results of a study which examined the administration of standardized tests in eight kindergartens. It concluded that "testing practices in five of the eight kindergartens were so nonstandardized as to render their test scores incomparable and quite possibly unreliable as well." It concludes that using results from such administrations is likely to lead to unsound educational decisions.

4. *THE IMPACT OF THE MISUSE OF TESTS BY THE PUBLIC SCHOOLS*

Madaus, G.

"Test Scores as Administrative Mechanisms in Educational Policy", *Phi Delta Kappan* (May 1985), p. 611 - 618.

This article provides a very brief history of testing in American public schools. It suggests reasons for the increased use of tests in the schools. It describes various problems with test use including: a loss of authority for teachers' professional judgment; a loss of local control over education; narrowing the educational curriculum; teaching to the test; a disproportionate impact on disadvantaged students; and the elevation of coaching over teaching.

(This document is available from NCAS, 100 Boylston St., Boston, MA 02116.)

Pullin, D.

"Educational Testing: Impact on Children At Risk". *NCAS Backgrounder* (December 1985).

This report discusses: the increased use of standardized tests in the public schools; their use as barriers to educational opportunity for at-risk children; the misclassification of minority students; the impact of tests on handicapped students; and the general measurement limitations of tests.

(This document is available from NCAS, 100 Boylston St., Boston, MA 02116.)

A. *Impact on Educational Goals and Curriculum*

Dorr-Bremme, D.W. and Herman, Joan L.

***Assessing Student Achievement: A Profile of Classroom Practices.* (Center for the Study of Evaluation: Los Angeles, 1986).**

This article reports on a nationwide survey of teachers and principals in 114 school districts. It seeks to identify the amount of time devoted to testing, using test results, teachers' and principals' perceptions about testing, and issues of equity as a result of standardized test use. Concerns about impact of standardized test use include the fact that most teachers pay more attention to standardized test results for low SES students than for high SES students, the tests can reduce time spent on other curricular goals as tests can lead to narrowed curriculum. The authors call for a more "rational" relationship between teacher-designed tests, external texts and the curriculum, but emphasize that the curriculum must drive the tests and not vice versa.

Meier, D. **"Why Reading Tests Don't Test Reading".** ***Dissent*** **(Winter 1982-83).**

Reading instruction aimed at increasing standardized test scores is shown to hinder learning to read, since higher test scores do not necessarily indicate improved reading ability. The author concludes that reading and other tests are biased against minority and low-income youth. Tests cannot measure the utility or effect of school reforms: ". . . testing not only fails to be helpful but sabotages good education."

National Academy of Education. ***The Nation's Report Card: Improving the Assessment of Student Achievement,*** **Review of the Alexander/James Study Group Report (National Academy of Education, 1987), p. 47.**

Although this review discusses the recommendations of the Alexander/James Study Group to dramatically expand the National Assessment of Educational Progress (NAEP), it also provides some general discussions of the problems of testing. In particular, it notes how testing can narrow the curriculum, threaten educational equity, and undermine educational goals that are not easily quantifiable.

(This report is available from NAE, Harvard Graduate School of Education, Longfellow Hall, Cambridge, MA 02138.)

Salganik, L.H. **"Why Testing Reforms Are So Popular and How They Are Changing Education".** ***Phi Delta Kappan,*** **(May 1985), pp. 607 - 610.**

This article links the growing use of standardized tests in the schools with a loss of public confidence in teachers. Because increased reliance on testing undermines the authority of teachers' judgments, a cycle of declining professional authority and declining public confidence is created. It also notes that policy issues, such as educational equity, the goals of schooling and control over school decisions are masked by the emphasis on technical questions regarding testing.

Steelman, L.C. & B. Powell. **"Appraising the Implications of the SAT for Educational Policy".** ***Phi Delta Kappan,*** **(May 1985), pp. 603 - 606.**

This article focuses on the misuse and misinterpretation of state averages for SAT scores. A brief discussion concludes that overreliance on standardized tests has led to teaching to the test and results in school staff discouraging potentially low scoring students from even taking the test.

Suhor, C. **"Objective Tests and Writing Samples: How Do They Affect Instruction in Composition?" Phi Delta Kappan, (May 1985), pp. 635 - 639.**

This article criticizes the use of multiple-choice standardized writing tests because they undercut "real" writing instruction. The author presents a case for the use of computerized "writing sample" tests as an alternative.

Tyson-Bernstein, H. ***A Conspiracy of Good Intentions: America's Textbook Fiasco.*** **(Council for Basic Education: Washington, D.C., 1988).**

This report discusses the quality of textbooks in the public schools. It concludes that the increasing emphasis on testing has been a major contributor to the declining quality of American textbooks. In particular, it indicts the current "curriculum alignment" movement which affects school districts in 22 states. It encourages the use of a more diverse curriculum and set of student assessment mechanisms.

(This report is available from the Council for Basic Education, 725 15th St., N.W., Washington, D.C. 20005.)

B. *Impact on Student Progress and Achievement*

Raudenbush, S. **"Magnitude of Teacher Expectancy Effects on Pupil IQ as a Function of the Credibility of Explanatory Induction — A Syntheses of Findings from 18 Experiments", Journal of Educational Psychology (January 1984), pp. 85 - 97..**

This article reanalyzes the "Pygmalion Effect", originally identified and described by Robert Rosenthal and Lenore Jackson in 1965 in *Pygmalion in the Classroom*. In examining data from 18 different experiments, it discovered that the Effect was most prevalent among students entering seventh grade, but not those in grades three to six. It points out that the effect is most likely to occur when teachers know little about their students beyond the statistical information they are provided. Where that information (including test scores) under- or overestimates students abilities, teacher behavior affects student achievement (regardless of previous student achievement levels.)

Willie, C.V. **"The Problem of Standardized Testing in a Free and Pluralistic Society". *Phi Delta Kappan,* (May 1985), pp. 626 - 628.**

This article discusses the disproportionate impact test use has on minority students. It notes that tests ignore the diversity in "mental patterns" among various American ethnic groups and implicitly undercuts their value and legitimacy.

C. *Impact on Local Control*

Wise, A.E. **"Legislated Learning Revisited", *Phi Delta Kappan* (January 1988), pp. 328 - 333.**

The article demonstrates how the increasing use of standardized testing undermines local control over the public schools. The author describes various ways in which test use increases state and national control over education. He also discusses other impacts of increased test use such as narrowed curriculum, loss of teaching time, lower

Subscribe to FairTest Publications

FairTest publications will keep you up-to-date on the growing movement for fair & open testing. You will be able to do more than just read about FairTest's campaigns—you can help make them successful!

_____ *FairTest Examiner*, published quarterly...Individuals $15/year; Institutional checks $25

_____ *Fallout from the Testing Explosion: How 100 Million Standardized Exams Undermine Equity and Excellence in America's Public Schools*, by Noe Medina and Monty Neill with FairTest staff (45 pages)...$8.95

_____ *Standardized Testing Reform Sourcebook, 1988* (200 pages)...$25.
(Currently out of print, available Summer, 1988)

_____ *Sex Bias in College Admissions Tests: Why Women Lose Out*,
by Phyllis Rosser and FairTest staff, 2nd edition (40 pages)...$7.95

_____ *The Reign of ETS*, by Allan Nairn and Ralph Nader (550 pages)...$30

_____ *Beyond Standardized Tests: Admissions Alternatives that Work*,
by Amy Allina with FairTest staff (15 pages)...$5.50

_____ *None of the Above: Behind the Myth of Scholastic Aptitude*, by David Owen (327 pp. lists at $7.95)...$5.95.

Updates, issued bi-annually (spring & fall). $7/year/topic:

_____ *Testing Reform Initiatives*
_____ *K-12th Grade Testing*
_____ *University Admissions Tests*
_____ *Teacher Testing*
_____ *Employment Testing*
_____ *Sex Bias in Testing*
_____ *Test Coaching*

Order total _____

Tax-deductible contribution _____

TOTAL ENCLOSED _____
(make check payable to FairTest)

_____ Back issues of *FairTest Examiner* $4 each: Vol___, No.___ (all 5 for $18.00)

_____**Moving.** I have moved/will be moving. My new address is (Please enclose old label or write old address.):

Name _____ Title_____

Organization_____Phone No. _____

Address _____Zip_____

Make checks payable to FairTest and mail to: FairTest, P.O. Box 1272, Harvard Square Station, Cambridge, MA 02238

ERIC

FairTest

National Center for Fair & Open Testing

Fallout From the Testing Explosion

More than one hundred million standardized, multiple-choice tests were administered to America's public school students during the 1986-87 school year, according to a new study by FairTest, the National Center for Fair and Open Testing. The report finds that over-reliance on standardized exams damages equal opportunity and school quality and "will lead to a worse, not better, public understanding of the schools and a weaker, not a stronger, educational system."

Fallout From the Testing Explosion: How 100 Million Standardized Exams Undermine Equity and Excellence in America's Public Schools is based on a comprehensive survey of state education administrators and a sample of local school officials, plus a review of research literature on tests and their use.

FairTest researchers found bias against minorities and low income children, as well as serious problems in test construction, reliability, validity and administration. They conclude that use of these instruments leads to misplacement of students, narrowing of pedagogy and curriculum, and loss of control over education to the unregulated testing industry.

This report should be useful for teachers, administrators, parents, policy makers, and legislators concerned about the growing use of tests in our schools. An extensive annotated bibliography makes this report a particularly valuable resource.

To order Fallout From the Testing Explosion, send $8.9: (postpaid) to: FairTest, P.O. Box 1272, Cambridge, MA 02238.

The FairTest study found that Southern states and larger school districts test most often. Most states with no mandated tests have small minority populations. Newark, New Jersey, reported the highest rate of testing, with an average of nine standardized tests given to each student during 1986-1987.

FairTest also found that least 13 states and the District of Columbia are "aligning" their curricula to reflect the content of standardized tests. "Many schools have embarked on a single-minded quest for higher test scores even though this severely narrows their curriculum," said Dr. Neill.

"Standardized tests have become the all-powerful gatekeepers of American education," FairTest concluded. To counteract the overuse misuse of standardized tests, the FairTest report advocates four reforms:

- All tests must be constructed, validated and administered so that they measure pertinent, not extraneous, differences between students;

- Tests should be open so that publishers claims can be independently investigated;

- No test score should be used as the sole or primary factor in any "high stakes" educational decision;

- New assessment instruments must be developed as alternatives and supplements to standardized multiple-choice tests.

To advance its agenda, FairTest is sponsoring a national testing reform conference on "Breaking Down the Barriers" in Washington D.C. on Jan 17 and 18. Leaders of civil rights, education reform, and consumer advocacy organizations are expected to attend a Friday afternoon Capitol Hill reception featuring a speech by Sen. Albert Gore (D.-Tenn) and Saturday workshops and seminars at Howard University Law School.

Copies of "Fallout From the Testing Explosion" are available for $8.95 from FairTest, P.O. Box 1272, Cambridge, MA 02238.

FairTest is a national non-profit research and advocacy organization funded by grants from the Ford Foundation, Rockefeller Family Associates, and individual donors.

- - 3 0 - -

attached: Key Findings from "Fallout from the Testing Explosion"
State-by-State chart of required tests

Copies of the full report are available on request