PART I
Advancing the Conversation

Oppression, Privilege, and High-Stakes Testing

Carl A. Grant
Department of Curriculum and Instruction
University Wisconsin—Madison

Performance measures and educational tests are not really about measurement. They are about political communication. (Kentl, 1998, p. 47)

Today, discussions about high-stakes testing of students rage in the popular media, in the Oval Office, in Congress, and within state houses across the nation. Also, reactions to these discussions of high-stakes testing are underway in schools and communities, among and between teachers, students, parents, and administrators.

Former President Bill Clinton is credited with initiating the present discussion of high-stakes testing. In his 1997 State of the Union address, Clinton challenged the nation to undertake “a national crusade for education standards—not federal government standards, but national standards—representing what all students must know to succeed in the knowledge economy of the twenty-first century” (National Research Council, 1999, p. 14). Clinton argued that “Every state should adopt high national standards, and by 1999, every state should test every fourth-grader in reading and every eighth-grader in math to make sure these standards are met” (p. 14). Clinton declared, “Good tests will show us who needs help, what changes in teaching to make, and which schools need to improve” (p. 14). With these words, the focus on standards and high-stakes testing increased tenfold.

Purpose

This article discusses the oppression and privilege of high-stakes testing. I discuss who is oppressed and who is privileged because of high-stakes testing. The oppression I speak of is both mental and physical oppression. Oppression, according to the American Heritage Dictionary (1985) is “a feeling of being heavily weighted down, either mentally or physically” (p. 872). High-stakes tests cause a great deal of mental anxiety for students, teachers, and parents. The results of high-stakes tests often determine the future location (school, job, home) and the resources and status of those taking the tests. Privilege, according to the American Heritage Dictionary (1985) is “a special advantage, immunity, permission, right, or benefit granted to or enjoyed by an individual, class, or cast” (p. 986). Privilege associated with high-stakes tests is concerned with the privilege that politicians, test publishers, and others derive from using high-stakes tests as a way to garner political muscle and financial gain. I discuss this point later in the article.

I write this article wearing two hats: one as a scholar whose research interests are issues of equity, power, and education that is multicultural, and the other as a grandparent whose grandchildren are caught up in the high-stakes test reform effort, even though one, Gavin, is only in first grade and the other, Amaya, is in day care. Before I proceed with my discussion of oppression, privilege, and high-stakes testing, I begin with a defini-
tion of high-stakes testing and a brief overview of the history of tests in education. There are more comprehensive historical accounts of tests and testing in the educational literature (Hanson, 1993). My contribution here is to prick our collective memory and sharpen our vigilance, because tests and testing have caused the oppression of students of color over the decades.

Definition

There is agreement in the educational literature and popular press that high-stakes means consequences; that is, rewards or sanctions accrue to some combination of students, teachers, principals, and schools as a result of how well (or poorly) students perform on tests (Secada, 2003). Educational literature also reports that high-stakes tests extend beyond the concept of standardized testing to denote the attaching of high-stakes consequences to particular educational activities such as ability grouping, retention, promotion, or graduation (Huebert & Hauser, 1999); and to serve educational policy goals, such as setting high standards for student learning and raising student achievement (National Research Council, 1999, p. 2). In addition, the educational literature claims high-stakes assessment often uses the results of a single measure to make important decisions about student progress or the effectiveness of instruction (McLaughlin, 2002).

These and most other definitions of high-stakes tests are clear and sterile in expressing meaning, consequence, unit of measurement (single), and relation to educational policy. This is, it can be argued, as it should be. However, because our educational system has a checkered past in providing equality and equity to students of color, my question is whether clarity and sterility are enough. I believe they are not. The impact of power and privilege, and attention to race, ethnicity, gender, and English language learners also need to be contextually added and foregrounded in any understanding and analysis of high-stakes tests. Discussions of tests and testing that do not include how race, gender, class, and language influence the discourse result in biased understandings.

History

Tests have been a part of Western civilization for centuries. For example Plato tested the young slave Anytus about his knowledge of mathematics. However, in the early 1900s, social scientists and educators increasingly became interested in the science (the quantitative and measurable aspects) of education. Educators wanted to know about the general intelligence of students. In 1905, Binet and Simon published a scale to measure intelligence. The scale was revised in 1916 by Terman and became known as the Stanford Revision of the Binet scale, or the IQ test. The Terman scale was an answer to social scientists’ and educators’ demands for a tool to ascertain the general intelligence (IQ) of students. The purpose of these tests, Hanson (1993) argued, is to measure not so much what the student has already learned, but how much the student is likely to learn in the future, or better, the student’s ability to learn.

During this same time, teachers, principals, and parents wanted to measure pupils’ achievement of subject matter. This was, in part, because of European immigration. Tyack (1974) told us that with the increased immigration population came the demand for more efficient management, including “the objective and efficient classification, or grading of pupils” (p. 44).

Using tests to classify and grade students began in the 19th century, claimed Mulhern (1946). He argued that “Educational, as distinct from psychological, measurement began when J. M. Rice gave a spelling test to 30, 000 children in 1897” (p. 423). Schools wanted to know how well students were learning to spell. From this beginning, Mulhern (1946) contended, “came our standardized tests in school subjects” (p. 423). He claimed in “1908, C. W. Stone published arithmetic tests; in 1910 Thorndike developed a handwriting scale; and in 1912, the Hillegas developed a composition scale” (p. 423). Early discussions about how these “objective” data were used to best educate African American children and other students of color, many of whom were in poor and segregated schools, are absent from the literature I reviewed. At best, what one discovers are discussions about how poorly Blacks and lower class students perform on IQ tests (Hanson, 1993).

The importance of tests as ways to identify achievement and intellectual ability grew during World War I and mushroomed during World War II. Uncle Sam wanted to know, who—and let me add that the who I am speaking of are White men—in the armed forces had the potential to become officer candidates (Lemann, 1999), and who could be assigned to the more intellectually challenging tasks (e.g., flying or navigating) of fighting a war. Army Alpha and Beta tests, as they were called, tested the IQ of recruits. Lemann (1999) stated, “Every Army inductee, ten million men in all before the end of the war (World War II), took an adapted IQ test called the Army General Classification Test, and was assigned in part on the basis of the result” (p. 53). These tests mostly assessed a soldier’s social class and sense of monetary resources. They asked such things as where Yale University is and what kinds of items could one purchase in certain stores.
In 1957, the demand to know how well students were achieving intensified when the Soviet Union launched Sputnik. The American public was shocked and dismayed that the Soviets were leading the space race. Kliebard (1995) reported that within days after the launching of Sputnik, the mass media had settled on a reason for the Soviet technological success. Schools in the United States were identified as the problem and became the whipping post. Americans were irate and demanded to know why “Nikita” and “Ivan” achieved higher scores in science, mathematics, and reading than “Dick” and “Jane.”

It is important not to forget that during this time the United States was dealing with another educational problem: school desegregation. However, most of the educational literature on testing that I reviewed does not relate the inferior and segregated education African American students receive to the Sputnik debate. No critique existed about the contribution African American students could provide if they were not attending poor, segregated schools. It was as if Black students were not considered worthy enough or they were not considered to have the academic ability to make them a viable asset in this race.

The criticism of schools that the U.S. public heard was that the life adjustment curriculum schools were teaching was too soft. Critics argued that while students in the United States were learning to get along with one another or learning how to bake a cherry pie, Soviet students were being steeped in the hard sciences and mathematics needed to win the technological race (Kliebard, 1995). Vice Admiral Rickover, a staunch critic of the educational system, made a statement that captures the tenor of the time: “The school’s concern is with the intellect alone” (cited in Kliebard, 1995, p. 227). Such a statement ignored the social engineering (school desegregation) that schools were conducting and supported a climate of testing.

In 1958, a year after the launching of Sputnik, Congress responded to critics of education and proponents of testing with the passage of the National Defense Education Act (NDEA). The intent of the act is clearly stated in the last sentence of the first paragraph: “The defense of this Nation depends upon the mastery of modern techniques developed from complex scientific principles” (Kliebard, 1995, p. 227). To promote the defense of the nation, the NDEA ushered in curriculum changes in science, mathematics, and foreign language. Also, the NDEA focused attention on strengthening schools’ guidance services to identify academically gifted students. Thus, during the time between World War II and the NDEA and motivated by the launching of Sputnik, the idea of national testing gained a firm foothold in the educational landscape.

Supporting the idea of national testing was the U.S. public, who feared the Russians were coming and demanded major educational reform. Further, many in the United States were coming to believe in a new economy that rewarded those who were highly educated and had technology skills and punished those who had a minimum of education and few technology skills. To support the high value the new economy placed on educational attainment and job skills, the U.S. public demanded that the education system make sure that children had the knowledge to be productive citizens (Sacks, 1999). A companion of the new economy was an ideology, social Darwinism, that did not support federalism for the poor. Mental measurement tools (intelligence, ability, personality, and academic achievement tests) served as the gatekeepers to the new economy (Sacks, 1999).

Also, serving to strengthen the idea of national testing were the efforts of American College Testing (ACT) and Educational Testing Services (ETS). Both ACT and ETS contributed to the argument for testing when they promoted the benefits of tests to high school seniors and the colleges they applied to by informing them of the student’s “scholastic aptitude,” a close relative to IQ. The hope of Henry Chauncey, the president of ETS, was that the Scholastic Aptitude Test (SAT) would affect the life course of practically everybody in the United States (Lemann, 1999).

Critics of the proposed national assessment program predicted damaging social consequences to education and argued against it. Hand (1965) contended that the results would be a rigid national testing program similar to that used by ministries of education in European, African, and Asian countries. The Association for Supervision and Curriculum Development claimed that national assessment would be a step toward federal control of education and a uniform national curriculum (Van Til, 1974). These and other responses against national assessment were not enough to halt this surging idea. Science and measurement was winning the day. National testing was not put into practice, but the idea was strongly embedded in the minds of many in the United States. National testing was like an actor waiting in the wings for the cue to take the stage and steal the show.

A major step toward national assessment occurred in 1963 with the development of the National Assessment of Educational Progress (NAEP). The NAEP’s purpose, supported by federal dollars, is to be a national representative assessment of what students in the United States know and can do in several content areas. The first administration of the NAEP in 1969 tested students’ science knowledge; in
1970, reading and literature were tested, and writing was assessed in 1973 (Salinger, 2002). After the establishment of the NAEP, the federal government support of national or large-scale testing grew. Large-scale testing was a part of the 1965 Elementary and Secondary Education Act legislation. Here, policymakers, politicians, and academicians were concerned with the achievement of poor and Black students enrolled in social action (compensatory) programs, such as Head Start and Follow Through. Large samples of children (e.g., 1, 875) from different sites (e.g., Lee County, Alabama; Portland; St. Louis; Trenton, New Jersey) were selected for study to answer such questions as these: "What are the components of early education that are associated with the cognitive, personal, and social development of disadvantaged children? What are the environmental and background variables that moderate these associations?" (Shipman, 1973, p. 148).

"Nevertheless, high-stakes tests are oppressive for the following reasons: (a) high-stakes tests undermine quality teaching and learning, and (b) high-stakes tests make students vulnerable in the classroom to a narrowly focused curriculum in which teachers teach to the test."

The testing of children in compensatory programs, however, was flawed and oppressive, because the operating definition of "compensatory" contended that because the child comes from a poor home, certain generalized supports for the kind of development that are deemed important in school are probably missing or deficient. Therefore, the compensatory thesis argues that is important to provide children support through some means other than the family (Anderson, 1973). Although it might be true that students from poor homes need help, help needs to be more broadly conceptualized to include help that would change school policy and practice to better meet the needs and interests of the students instead of assuring that students conform to the demands of the school.

States, too, promoted large-scale testing. In the 1970s, the minimum competency testing movement, which used large-scale, standardized achievement tests to assess student achievement (and school accountability), grew in popularity. Some states (e.g., Florida) implemented minimum competency testing as a partial requirement for high school graduation. By the mid-1980s, 33 states had mandated some form of minimum competency testing. By the mid-1990s, somewhat in response to the National Commission on Excellence in Education (1993) report, A Nation at Risk, which claimed that there was "a rising tide of mediocrity" (p. 5) in U.S. public education, and in response to the growing standards movement, 18 states had test-based requirements for high school graduation. Also by this time, in about 25 states, students' performance had serious consequences for their schools, including funding gains or losses. School losses included autonomy or accreditation, and even the threat of external takeover (National Research Council, 1999). It was at this same time, as noted earlier, that former President Clinton increased the rhetoric of standards, which was shortly followed by high-stakes tests.

Oppression

So far my argument can be interpreted as contending that reasonable and wise people promoted tests to save the day for the United States. The Russians did not come; schools began teaching more reading, mathematics, science, and computer skills; and the business world had a larger pool of qualified applicants. So, one might ask, "What's the problem with testing?"

Accountability is important and necessary in our society. Research data support the positive effects of testing on students and teachers (Raymond & Hanushek, 2003). Additionally, I accept Secada's (2003) claim that, "in some cases high-stakes testing does help focus attention on what is important" (p. 6). Nevertheless, high-stakes tests are oppressive for the following reasons: (a) high-stakes tests undermine quality teaching and learning, and (b) high-stakes tests make students vulnerable in the classroom to a narrowly focused curriculum in which teachers teach to the test.

High-Stakes Tests Undermine Quality Teaching

High-stakes tests are often multiple-choice, fill-in-the-bubble or square with a number 2 pencil tests. As such they undermine quality teaching and learning because they are imperfect instruments for assessing a number of activities and exercises that students experience. In instructionally engaging schools, students create interdisciplinary projects, solve real-world problems us-
ing multiple solutions, synthesize and evaluate multiple forms of data and information from a variety of sources, and apply their learning to new situations. Such engaging instruction is not easily assessed by a fill-in-the-bubble examination. Popham (2001) agreed when he stated that high-stakes tests undermine quality teaching because of a mismatch between what is taught and what is assessed, and these mismatches are often unrecognized or ignored. Also, Frederiksen (1984) stated, “The real test bias is that multiple-choice tests tend not to measure the more complex cognitive abilities” (p. 193). Additionally, Meier (cited in Neil, 2003) argued that “multiple-choice reading tests are not described and used as measuring a few limited aspects of ‘reading skills’; they are erroneously described as measuring ‘reading’” (p. 5).

Furthermore, teacher quality is undermined when teachers are anxious and concerned about the impact of high-stakes tests. Davis, Davis, and Leppo (1999) conducted a study of Sailie Mae’s First Class Teacher program. The survey study interviewed teachers who exhibited outstanding performances in their first year of teaching in kindergarten through Grade 12 to discover their ideas on how to understand and improve elementary and secondary education. The First Class Teachers had much to say about the nuts and bolts of education, including ideas about high-stakes testing. They argued that local, state, and national tests restrict their time to work with students. Also, they claimed teaching subject content beyond the subject areas being tested has become much harder and more frustrating. They reported that the publication of “school report cards” that identify schools that are meeting the set standards as well as those schools that are not meeting the expected performance standards is having a negative impact on teachers as well as students and parents.

The First Class Teachers contended that as they are subjected to the rising emphasis on attaining expected test scores, they are compromising their own beliefs about their roles as teachers, and the importance of meeting students’ needs at their own level of comprehension (Davis et al., 1999). In addition, First Class Teachers claimed they are frustrated with trying to adapt their teaching methodology to new assessment expectations when, at the same time, they are not given adequate professional development and resources. A real fear, they argued, is that they are going to succumb and just teach to the test.

Finally, teaching quality is undermined because high-stakes tests interfere with making learning exciting and culturally relevant. High-stakes tests, because they demand test-preparation attention, promote ways of teaching that are often boring and neglectful of problems and issues concerned with race, class, gender, and sexuality. Such neglect is an enactment of power that circulates in oppressive ways and is antithetical to culturally relevant teaching, which Ladson-Billings (1994) stated is “an approach to teaching and learning that empowers students intellectually, socially, emotionally, and politically by using cultural referents to impart knowledge, skills and attitudes” (p. 62).

The American Evaluation Association (AEA, n.d.; see www.eval.org) reported how high-stakes testing programs disrupt culturally relevant teaching. AEA contended that high-stakes testing channels educational offerings to satisfy narrow, test-defined state standards, rather than addressing the differential needs of students in local schools. Schools, the AEA claimed, are drawn into narrow conceptions of teaching and education that leave children deprived of the history, cultural perspective, personal experience, and
interdisciplinary nature of subject matter. Further, the AEA contends high-stakes testing assumes that all children, including English language learners and special education students, learn in the same way at the same rate, and that they all demonstrate their achievement equally on standardized tests.

Students Are Academically Vulnerable Due to High-Stakes Testing Reform

Valenzuela (2002) argued that the large backdrop of inequality at both the societal/institutional and personal level makes U.S. Mexican students vulnerable in a high-stakes testing environment. This vulnerability receives, at best, a cursory and narrow discussion in the high-stakes testing discourse. Valenzuela contended that several factors contribute to students’ academic vulnerability: cultural insensitivity in schools, segregation of students, unqualified teachers, and poorly educated parents.

Schools, Valenzuela (2002) posited, are culturally subtractive to U.S. Mexican students. “Historically, rather than building on children’s social, cultural, and linguistic competencies, schooling, as a tool of Americanization, has played the role of subtracting from children their language, culture and community-based identities” (p. 3). Such subtractions, Valenzuela argued, include narrowing the curriculum by teaching to the test, marginalizing students’ language and culture in the curriculum, and retaining students in grades or relegating them to test-exempt status categories to produce positive test results and school ratings.

Racial segregation of schools, Valenzuela (2002) claimed, is a major factor that contributes to poor academic success. Quoting Orfield (1992), she noted that “Hispanic students attending schools in California and Texas experience greater segregation than Blacks in Alabama and Mississippi” (p. 3). Segregation is correlated with underachievement on numerous indexes, including standardized tests, high dropout and retention rates, and a low percentage of students entering and finishing college. In addition, Valenzuela observed that African American and Latino students disproportionately attend the lowest accredited schools, and she argued that one in five public school students in Texas are taught by teachers who are not certified in the subjects that they teach. The use of uncertified teachers increases students’ vulnerability for high-stakes testing as well as other academic challenges students face in later life, such as applying to college or seeking employment opportunities.

Complicating and adding to the vulnerability of students, Valenzuela (2002) contended, is the lack of education that the parents of first-generation Mexican immigrant students have received. Parents have attained an average of 6 years of schooling, and third-generation adult Mexican Americans complete an average of 9.3 years of schooling, with a dropout rate of 56%. This lack of schooling contributes to Mexican American parents’ lack of assertiveness to challenge their children’s failure to pass the Texas Assessment of Knowledge and Skills (TAKS) exams before the grade placement committees that decide on the promotion or retention of students. Valenzuela (2002) stated:

The student’s parents or guardians may appeal the decision to a “grade placement committee” comprised of themselves, the principal (or designee), and the child’s teacher corresponding to the subject area on the test that they failed. However, not all parents will appeal, leaving the decision intact. Among those that do, the process is prejudiced against them, particularly if they are poor, non-English-speaking, or minority. (p. 10)

Johnson and Johnson (2002) provided a revealing account of how students are vulnerable and oppressed during this high-stakes test reform. Johnson and Johnson spent one school year as full-time teachers in a rural school in Louisiana. Redbud Elementary has no playground, no library, no hot water, a shortage of teachers, and hardly any paper supplies. However, Redbud has a good deal of unkempt and unclean school property, including rats and roaches. Johnson and Johnson reported that most of their students lived in single-parent homes, and some of the homes were without electricity, running water, or floors. In this climate of poverty and despair, the state is rigorously monitoring mandated high-stakes tests, in lieu of other student and school needs.

Johnson and Johnson (2002) described the students and their perspectives in the following two statements. The first one is from a time shortly after students received their test results.

Sounds of fourth graders crying carry down to where my third graders sit. “Why are all those kids crying?” asks Jaron. “They flunked,” says Kenziah softly. I walk down to the fourth-grade classrooms. I try to console inconsolable children. The pitiful scene is too much for me. I walk back to my third graders. (p. 178)

The following statement of anger and frustrations by Johnson and Johnson (2002) is indicative of similar feelings by numerous teachers across the country.

I am filled with anger at the state bureaucrats and politicians who designed and mandated this uncompromising accountability system. Our Redbud pupils have so many strikes against them. They often are sick. They have rotten teeth and cry because of severe toothaches. Many come from dysfunctional homes. They are ill-clad and wear ill-fitting shoes. Several do not eat enough food or enough rest. They live with acute poverty in substandard
homes, often surrounded by drug dealers and users and drunks. The harsh accountability system imposed by the state kicks them further. I am not opposed to testing. Well-designed tests can give educators useful information. Ideally, the results would inform districts about needs for remediation. Tests should be used to enlighten, not to torment. (p. 176)

Although some—teachers and students—live in frustration because of high-stakes tests, there are others who receive privileges because of high-stakes tests.

Privilege

There are at least three kinds of privileges associated with high-stakes testing: the revenue that test developers receive, the leverage and favoritism politicians gain from promoting high-stakes testing, and the advantage students from higher socioeconomic status groups have when they take high-stakes tests. These first two are described together as revenue and political gain.

Revenue and Political Gain

The amount of revenue from states and federal education budgets directed toward tests has greatly increased since the accountability reform effort received prominent attention. Sacks (1999) stated that “between 1982 and 1994, standardized test sales grew faster than sales for school and college textbooks, mass market paperbacks, book clubs, and other segments of publishing” (p. 6). Sacks added, “Americans are taking as many as 600,000 standardized tests each year in schools, colleges and universities, and the workplace” (p. 6). Metcalf (2002) contended that over the past 5 years, the money spent on state testing programs more than doubled from $141 million to $390 million. Hoff (2001) claimed that in 2001, states were spending almost a half-million dollars a year on testing and were bracing for a substantial increase with the No Child Left Behind Act.

The No Child Left Behind legislation, Karp (2002) observed, appropriates about $400 million each year for the next 6 years to develop new tests, and he argued that according to a report in Time magazine, the figure will be more like $7 billion if each of the 50 states is to implement a high-quality testing program. Johnson and Johnson (2002) contended that the state of Louisiana would spend more than $56 million on its accountability program for 2001–2002. They stated the following:

We can’t begin to estimate the total cost being paid by school districts for expensive electronic and print materials purchased in the hopes that students will get higher test scores. This troubles us in a school that has no hot water for the children, no library, no playground equipment, no art classes, a shortage of textbooks for each child, and uncertified teachers in some of the classrooms. (p. 154)

Who are the recipients of this financial windfall? Is it just good capitalism? Metcalf (2002) argued that a small number of companies are in a position to make enormous sums of money from high-stakes tests. Metcalf reported that the three largest textbook publishing companies—McGraw-Hill, Houghton-Mifflin, and Harcourt General—are in position to harvest the windfall. They are assertively making certain that they have a voice in education decisions by having very active lobbyists on Capitol Hill. Metcalf also pointed out the close ties between the Bush family and McGraw-Hill. He described a close relationship that began in the 1930s and continues today:

The amount of cross-pollination and mutual admiration between the Administration and the empire is striking: Harold McGraw Jr. sits on the national grant advisory and founding board of the Barbara Bush Foundation for Family Literacy. McGraw in turn received the highest literacy award from President Bush in the early 1990s for his contributions to the cause of literacy. Also, Harold McGraw III was selected as a member of President George W. Bush’s transition advisory team. (p. 4)

Also, Walsh (2001) argued that a Pearson executive noted that President Bush’s testing program “almost read like our business plan” (p. 8). In addition, Berliner and Biddle (1995) stated:

Some of those who accepted hostile myths about education have been genuinely worried about our schools, some have misunderstood evidence, some have been duped, and some have had other understandable reasons for their actions. But many of the myths seem also to have been told by powerful people who—despite their protestations—were pursuing a political agenda designed to weaken the nation’s public school, redistribute support for those schools so that privileged students are favored over needy students, or even abolish those schools together. (p. xii)

My point is not to wag a finger as if there is something illegal going on, but to call attention to the need to be more observant about all aspects of high-stakes testing, including who is making the money and gaining political leverage. Some are seeing education as a golden goose. For example, a recent interview with Friedman (2003) identified education as a growing industry. Friedman stated:

There will be a brand new industry: the education industry, a private, for-profit and nonprofit education industry. It will introduce competition in a way that’s never existed
before. And it’s a big industry. Total expenditures of elementary and secondary education in the United States are in the neighborhood of $300 billion. (p. 59)

Socioeconomic Advantage

Socioeconomic status and family background privileges some students over others, and the ability to do well on high-stakes tests is highly correlated with socioeconomic class. This thesis is fairly well documented in the educational literature. Sacks (1999) argued:

Although standardized tests have a relatively bleak record of predicting success in school and work, we know that they do not correlate exceedingly well with the income and education of one’s parents. Call it the “Volvo Effect.” The data is so strong in this regard that one could make a good guess about a child’s standardized test scores by simply looking at how many degrees her parents have and what kind of car they drive. (p. 8)

Ceci (1996) claimed that “Children from those families ranked in the highest social class usually achieve IQ scores about one standard deviation above the national mean, while children from the lowest ranking families usually achieve IQs about a half standard deviation below the national average” (p. 52). Ceci (1996) went on to claim that, “Social class exerts an influence throughout one’s entire history of schooling, and probably a vast number of low-income students do not even take SATs because of this. To the extent that this is true, the role of SES is even greater” (p. 54).

Conclusion

So, what has this scholar and grandfather of Gavin and Amaya learned from this analysis of the high-stakes testing literature? I have learned that during this high-stakes testing reform, the chances are diminishing of public schools richly contributing to their students becoming reflective, enlightened, and critical learners who have an appreciation and acceptance of social justice and global and national ethnic and racial diversity. If I wish for Gavin and Amaya to achieve a standard of education grounded in democratic principles, power sharing, and economic and political fairness it will be essential for their family members to take the leading role until public schools assume the primary responsibility for citizenship education.

Recently, a colleague I greatly admire and respect, who has two young children, told me that he was planning to home-school his children. I promptly replied, “Why?” My question had a ring of surprise and disappointment. You might ask why I was disappointed. Perhaps because I believe my colleague seemed to be giving up on public education, withdrawing his support and attention. I will not suggest to Gavin and Amaya’s mother that she should home-school them, but I do have a greater appreciation and respect for the statement made by my friend and colleague.

“I have learned that during this high-stakes testing reform, the chances are diminishing of public schools richly contributing to their students becoming reflective, enlightened, and critical learners who have an appreciation and acceptance of social justice and global and national ethnic and racial diversity.”

Finally, I learned that we must help the permanent underclass in our public schools. Let me explain. “Leave no child behind” is a catchy slogan for the high-stakes testing discourse. However, as Furr (2001), a reading specialist in a middle school in Charlottesville, Virginia, wrote in Education Week:

Americans generally underestimate, or perhaps choose to ignore, the range and depth of needs borne by children who fall behind in school. We devise interventions and targeted assistance—some of high quality and all well-intended—to help kids catch up or prevent them from falling through the cracks. But what such programs wind up doing, in effect, is obscuring the need for radical, systemic change. (p. 34)

Furr (2001) went on to tell readers about a child named Jevon, who he described as one of the 50 most “at-risk” fifth-graders at his school. Furr wrote that Jevon started school behind and has remained behind, although he was identified early as at risk. Jevon, he added, has gone to good public schools and has received excellent instruction and repeated reading interventions with bright teachers who care about him. However, progress in reading, according to Furr, is slow and incremental for Jevon. It does not show up on standardized tests, and it does not happen at the rate demanded by the public school system. Jevon’s case, according to Furr, is not extraordinary. There are many such students in his school and in schools across the nation. These students
are becoming what Furr rightly described as a “permanent underclass in our public schools” (p. 34).

Furr contended that many state legislatures have decided the best way to deal with the Jevons within the educational community is to flank them out. Furr (2001) stated that nearly “30 states now have some form of academic ‘standard of learning,’ accompanied by high-stakes tests that purport to measure whether teachers, schools, and students are doing their jobs well” (p. 34). By this he meant that all children are held accountable to the same standards and are tested at the same time using the same measures. No educator, he argued, and I agree, is opposed to standards of learning. However, high-stakes testing administered in this way guarantees the failure of Jevon and his underclass classmates in the school system.

In sum, high-stakes testing reduces teaching and learning for Jevon, Gavin, and Amaya to simple and routine procedures. It marginalizes their effort to learn and engage in a critical examination of themselves and society and detours their pursuit to become reflective and critical citizens.

References


