1. High-Stakes Testing: A Tool for White Supremacy for Over 100 Years

WAYNE AU

Educators first began applying standardized testing to school children as early as 1848, but it was not until the early 1900s that such testing developed into a widespread tool for measuring students, teachers, and people more generally in the United States. Given the ubiquity of testing in education policy and practice today, it is critical to understand that when we look at standardized testing historically and how it is used as the central mechanism for accountability within education policy, high-stakes, standardized testing has operated as tool for the maintenance of White supremacy in the United States.¹

In order to get to this understanding, in this chapter I first offer a basic overview of key concepts about high-stakes testing and then provide a brief, modern day history high-stakes testing in education policy. I then follow with how our modern day testing has roots in the racism of I.Q. testing and the eugenics movement around the turn of the 20th century and continue with how such testing was introduced into education in the United States, including the ways that these tests support White supremacy through the guise of offering individual inequality and “objective” measures of students. I then give an overview of the research evidence showing the disparate impacts of high-stakes testing on students of Color specifically, and I conclude this chapter with a brief description of the kinds of resistance building against high-stakes testing generally, as well as the White supremacists impulses embedded so deeply within the tests.

**Testing, The Basics**

In the most basic terms a standardized test is any test that is developed, administered, and scored in a systematic, standardized manner. There are two
general kinds of standardized tests: *norm referenced* and *criterion referenced*. A norm referenced standardized test compares one student’s test score to other students taking the same test, essentially comparing how individuals scored relative to the average across all test takers (the norm). A criterion referenced standardized test measures a student’s score against a set “criteria” of subject matter knowledge (Popham, 2001). Most standardized tests given in U.S. schools are now “high-stakes” tests in that students’ scores are used to determine rewards and sanctions for students, teachers, principals, schools, and school districts. Standardized tests are also high-stakes because many test scores are published and shared with the public, making students, teachers, principals, and schools feel the weight of public pressure regarding student test performance, regardless of the actual complexity behind the scores (McNeil, 2000).

Our modern era of education policy wrapped around high-stakes testing began with the 1983 publication of the Reagan Administration’s report on education, *A Nation at Risk: The Imperative for Education Reform* (National Commission on Excellence in Education, 1983). This alarmist report decried the poor quality of education in the United States. Even though much of the data used in *A Nation at Risk* was easily debunked (Berliner & Biddle, 1995), within a year of its release, 45 state level commissions on education were created and 26 states raised graduation requirements. By 1994, 45 states implemented statewide, standardized tests for grades K–5. U.S. President George H. W. Bush, in conjunction with state governors, developed the groundwork for his America 2000 plan—focusing on testing and establishing “world class standards” in schools. U.S. President Bill Clinton and Vice President Al Gore committed themselves to following through on George H. W. Bush’s plan, including a vocal commitment to pursuing a national examination system to meet new standards. By the year 2000, every U.S. state but Iowa administered a mandated, high-stakes, standardized test (Kornhaber & Orfield, 2001).

In 2001 President G.W. Bush advocated for federal education funding to be tied to test scores, and in 2002, the U.S. government passed the No Child Left Behind Act (NCLB) into law. NCLB notably required all students be tested in reading and math in grades 3–8 and once in high school, with future provisions that students be tested at least once at the elementary, middle, and high school levels in science. According to NCLB, if student standardized test scores did not meet stated goals for test score growth, schools faced penalties such as a loss of federal funding or the diversion of federal funds to pay for private tutoring, transportation costs, and other services (Karp, 2006). Upon taking office, President Obama selected Arne Duncan to lead the
Department of Education and implemented the federal Race to the Top program, which included monies for more testing as part of a broader education reform package promoting the use of tests to evaluate teachers, attacks on teachers unions’ right to collective bargaining, and the proliferation of charter schools (Kumashiro, 2012). Simultaneously, and as part of the Race To the Top agenda, national standards connected to high-stakes, standardized tests were developed as part of the Common Core State Standards as well (Karp, 2013, 2014).

Within the current context, high-stakes standardized tests function as the core for all education policy. From charter schools, to “value added measurement” and “student growth models” for teacher evaluation, to challenges to teacher tenure and collective bargaining, to school closings, to the use of undertrained and inexperienced Teach For America recruits in high needs and low needs schools, to the implementation of the Common Core national standards and tests, every current major educational initiative revolves around justifications associated with how different populations of students perform on high-stakes, standardized tests. These tests produce the data upon which important decisions about students, teachers, administrators, and schools are being made (Au, 2013). The centrality of high-stakes, standardized testing is the reason that it is important to understand their inherent racism and their role in enforcing White supremacy within the United States both historically and today.

A Brief History of High-Stakes Testing and White Supremacy in the United States

The origins of the use of standardized testing to assess and study large populations in the United States can be traced directly back to I.Q. testing and the eugenics movement of the early 1900s. Importing and distorting the work of French psychologist Alfred Binet, who developed the concept of “intelligence quotient,” or I.Q., U.S. psychologists distorted Binet’s original conception and use of I.Q. by injecting their own underlying presumptions about humans and human ability into their interpretations of test results (Au, 2009; Gould, 1996). For instance, as a psychologist and Army Colonel in charge of the mental testing of 1.75 million recruits during World War I, in 1917 Robert Yerkes worked with Henry Goddard, Lewis Terman, and others to develop the Alpha and Beta Army standardized tests to sort incoming soldiers according to their “mental fitness.” These cognitive scientists drew several dubious conclusions using this large pool of army recruit data, including that the intelligence of European immigrants could be judged according to their country
of origin: The darker peoples of eastern and southern Europe were less intelligent than their fairer-skinned, western and northern European counterparts. They also asserted that, based on their test results, African Americans were the least intelligent of all peoples (Giordano, 2005). As Karier (1972) explains:

Designing the Stanford-Binet intelligence test, Terman developed questions which were based on presumed progressive difficulty in performing tasks which he believed were necessary for achievement in ascending the hierarchical occupational structure. He then proceeded to find that according to the results of his tests the intelligence of different occupational classes fit his ascending hierarchy. It was little wonder that I.Q. reflected social class bias. It was, in fact, based on the social class order (pp. 163–164).

With the authority granted by the supposedly “scientific findings” of these psychologists, eugenicists—who believed in the genetic basis for behavioral and character traits they associated with gender, race, national and class difference—rallied around the idea that race mixing was spreading the alleged inferior intelligence genes of African Americans, other non-White peoples, and immigrants (Selden, 1999). In this way, these early standardized tests clearly operated as a tool in the maintenance of White supremacy in the United States because they provided supposedly “scientific” evidence to claims of White superiority.

Such standardized testing, which proved so useful as a technology for sorting people efficiently by race and class under the justification of pseudo-science, and a well established White supremacist socio-political order, quickly moved into public education in the United States. As Tyack (1974) explains:

Intelligence testing and other forms of measurement provided the technology for classifying children. Nature-nurture controversies might pepper the scientific periodicals and magazines of the intelligentsia, but schoolmen found IQ tests invaluable means of channeling children; by the very act of channeling pupils, they helped to make IQ prophecies self-fulfilling (p. 180).

Then a Stanford University professor of psychology, and under the sponsorship of the National Academy of Sciences, the above-mentioned Lewis Terman played a key role in adapting the above-mentioned army tests into the National Intelligence Tests for schoolchildren in 1919, and by 1920 over 400,000 copies of these tests had been sold nationwide. Terman and others also created the Stanford Achievement Test in 1922, and by late 1925, he reported sales of this test to be near 1.5 million copies. Surveys in 1925, 1926, and 1932 found that the majority of cities in the U.S. with populations over 10,000 people were using these tests to classify and sort students in to ability groups in their schools (Chapman, 1988; Haney, 1984). As Karier (1972) explains:
It was men like Thorndike, Terman and Goddard, supported by corporate wealth, who successfully persuaded administrators and lay school boards to classify and standardize the school’s curriculum with a differentiated track system based on ability and values of the corporate liberal society (p. 166).

Indeed, Karier’s (1972) analysis highlights a disturbing parallel for today’s context. Much in the same way that the Gates, Broad, and Walton Foundations, amongst others, influence education policy and reform in contemporary times (Barkan, 2011, 2012), Karier (1972) points out that:

From the very beginning of the century, the new philanthropic endeavors of corporate wealth were directed at influencing the course of educational policy. John D. Rockefeller’s General Education Board, which received its national charter in 1903, greatly influenced and shaped educational policy for Black American in the South, while the Carnegie Institute of Washington (1904) and the Carnegie Foundation for the Advancement of Teaching (1906) came to play a major role in shaping educational policy in both the South and the North. (p. 157)

These foundations were directly responsible for supporting the then-burgeoning movement for mass, standardized testing, and ironically the “values of the corporate and liberal society” to which Karier (1972, p. 166) refers, speaks to the ways that standardized testing was seen as a key to liberal notions of individual equality. These notions of individual equality and merit-based success based on hard work are deeply intertwined within the logics of standardized testing in the support of maintaining White supremacy, a topic I take up in the following section.

**Standardized Testing, Meritocracy and Whiteness**

In what follows, I trace some of the critical, foundational logics that undergird high-stakes, standardized testing and its conceptual link to White supremacy through the ideology of meritocracy and the idea of aptitude. I begin here with a discussion of the ideology of meritocracy and its connection to the presumed objectivity of the tests. Using the SAT as an example, I then address how, despite explicitly race-neutral processes for test development, racism and White supremacy are embedded into the structures of such tests. I follow by discussing the strong correlation between the structural inequalities (e.g., poverty) in the lives of students of Color with educational achievement, test scores included, and I conclude this section by explaining how the rational logics of standardized tests functionally require failure of student populations—particularly populations under the duress of structural racism and White supremacy in their daily lives.
One of the key assumptions undergirding the use of standardized tests to measure, sort and rank students is the idea that these tests are measuring students objectively and accurately—for if the tests are objective, then they truly are assessing the individual merits of students. In turn the individual students who have worked the hardest and who have the most merit, will rise to the top compared to their peers. This idea, that individual merit through hard work is what creates opportunity and success, is referred to as the idea of “meritocracy” (Lemann, 1999).

Despite producing clearly racist and classist outcomes, early standardized testing in the United States was viewed as providing a completely objective and value free measurement of human intelligence and ability (Au, 2009). This view of testing-as-objective-measurement has positioned these tests in a seemingly contradictory way: despite the historical and structural inequality associated with the tests’ outcomes, early advocates of standardized testing saw these tests as a means of challenging class privilege because they thought the tests were providing neutral, unbiased, and objective measures of individual effort and merit (Sacks, 1999). Under the assumption that standardized tests provide fair and objective measurement of individuals, such testing seemingly held the promise that every test taker is offered a fair and equal shot at educational, social, and economic achievement. The byproduct of this assumption was the masking of the racism and White supremacy that lay at the foundation of such testing at its outset. As Moore (2005) suggests,

In terms of race, we assume that who is tested, what is tested, and how tests are administered and interpreted have all been bathed by neutrality through “testing conditions” that include sterile classrooms, “expert” test givers, and the use of inanimate computers and “color blind” standards. (p. 184)

The ideology of meritocracy undergirding the use of standardized tests effectively conceal structural inequalities associated with racism and White supremacy under the cover of “naturally” occurring aptitude amongst individuals (Bisseret, 1979).

Vis-à-vis the ideology of meritocracy, the low achievement on standardized tests of non-Whites (as well as working class people and some immigrant groups) can then be simply and neatly attributed to the failure of individual students, individual groups, or individual cultures, and not attributed to existing structural inequalities that support White supremacy. Even the SAT, as Hartman (2007) explains,

...failed to stray from past definitions of what it meant to be American. Sure, after the implementation of the SAT Blacks and other peripheral groups were able to gaze upon increased opportunity. But this gaze was merely an apparition
for most. The SAT, consistent with other integration projects, did not accommodate Black or female identities—in this case, differing learning styles. In order for a more fully integrated society to emerge from the SAT, the onus was placed upon Blacks to accommodate to American identity. The SAT further entrenched a seemingly elusive White identity as the de facto American identity. (p. 143)

In this sense the standardized tests operate as a tool of White supremacy simply because they make racist outcomes of the tests appear as a product of the way the world works objectively and naturally—they “scientifically” justify the existing racial order, and they do so within a false promise of measuring individuals equally.

A more modern look at research on the SAT illustrates how the “science” behind standardized test development can functionally hide White supremacist outcomes. Kidder and Rosner’s (2002–2003) study of SAT questions finds significant White supremacist bias deep within the structure of test construction itself. In a study of over 300,000 SAT test takers, Kidder and Rosner examined the percentages of questions that Black, White, and Chicano (Mexican American) students answered correctly. Using calculations of average scores on each question by racial group, Kidder and Rosner determined what they called the “racial impact” of each test score question. For instance, if 50% of Whites and 30% of African Americans answered a particular SAT question correctly, the researchers assigned the question a 20% Black-White impact.

In their study, Kidder and Rosner (2002–2003) found that, “African Americans and Chicanos did not outperform Whites on any of the seventy-eight Verbal and sixty Math questions” (p. 148). Whites correctly answered 59.8% of the Verbal questions on average and African Americans correctly answered 46.4% of the Verbal on average, resulting in an overall 13.4% Black-White impact. Additionally Chicanos correctly answered Verbal test items at an average of 48.7%, giving an 11.1% Chicano-White impact. On Math questions Whites had an overall 58.4% correct answer rate, and African Americans had a 42% correct answer rate, giving an average disparate impact of 16.4%. Almost three out of ten Math questions averaged a 20% disparate impact between Whites and other groups. The average Chicano correct answer rate for Math questions was 46.5%, establishing an 11.9% Chicano-White impact.

Kidder and Rosner (2002–2003) find an explanation for these disparate test scores within the structure of the test design itself. The Educational Testing Service (ETS), who traditionally developed and administered the SAT at the time of the study, establishes statistically valid questions by using one section of the test as an experimental section, essentially trying out questions for possible use on future SATs. Based on the responses on the experimental
test items, the SAT designers then make decisions to either keep a question and use it in the regular sections of future tests or discard it as an unusable, “invalid” test item. Kidder and Rosner compared some of the regular test items with the experimental ones and found an important pattern. For example, on one Verbal test item of medium difficulty, 62% of Whites and 38% of African Americans answered it correctly (for a 24% disparate impact). This question was a test item from one of the regular, non-experimental test sections. By comparison, an item of similar difficulty used in the experimental test section resulted in African Americans outperforming White students by 8% (that is, 8% more African American students answered the question correctly than White students). Test designers determined that this question, where African Americans scored higher than whites, was invalid and was not included in future SATs. The reason for this was that the students who statistically on average score higher on the SAT did not answer this question correctly enough of the time, while those who statistically on average score lower on the SAT answered this question correctly too often. Rosner (2003) explains this process of psychometrically reinforced racism:

Each individual SAT question ETS chooses is required to parallel outcomes of the test overall. So, if high-scoring test takers—who are more likely to be white—tend to answer the question correctly in [experimental] pretesting, it’s a worthy SAT question; if not, it’s thrown out. Race and ethnicity are not considered explicitly, but racially disparate scores drive question selection, which in turn reproduces racially disparate test results in an internally reinforcing cycle. (p. 24)

At issue is the fact that statistically on average, White students outperform Black students on the SAT. Higher-scoring students, who statistically tend to be White, correctly answer SAT experimental test questions at higher rates than typically lower scoring students, who tend to be non-White, ensuring that the test question selection process itself has a self reinforcing, built-in racial bias. Couched in the language of statistical reliability and validity, the supposedly race-neutral process of test question development and determination ultimately structures in very race-biased results into the selection of the test questions themselves (Kidder & Rosner, 2002–2003), one that is arguably White supremacist in outcome regardless of the meritocratic impulses that in part defined the origin of the SAT. Further, it is worth noting that more recent research on the SAT has found other forms of racial and cultural bias built into the kinds of questions asked on the exam, where significant bias was found in the easier questions that appear earlier in the SAT test question sequence (Santelices & Wilson, 2010).

Similar to the SAT, other forms of modern day high-stakes, standardized testing have also proven not to be the paragons of meritocracy and have
instead proven to recreate race and class based inequality. As Berliner (2012) explains, test scores in the U.S. are more determined by structural conditions affecting students than individual effort:

Virtually every scholar of teaching and schooling knows that when the variance in student scores on achievement tests is examined along with the many potential factors that may have contributed to those test scores, school effects account for about 20% of the variation in achievement test scores…

On the other hand, out-of-school variables account for about 60% of the variance that can be accounted for in student achievement. In aggregate, such factors as family income; the neighborhood’s sense of collective efficacy, violence rate, and average income; medical and dental care available and used; level of food insecurity; number of moves a family makes over the course of a child’s school years;… provision of high-quality early education in the neighborhood; language spoken at home; and so forth, all substantially affect school achievement (n.p.).

While school does remain important, socio-economic factors outside of school—factors themselves that disproportionately impact communities of Color—simply have an overwhelming effect on educational achievement. However, this reality is effectively masked by the ideology of meritocracy embedded in high-stakes testing in the United States (Au, 2009).

The meritocratic assumptions of high-stakes testing in the U.S. are also belied by many of the logics that underpin the tests themselves. For instance, akin to systems of capitalist economics, systems of accountability built upon high-stakes, standardized testing cannot function if everyone is a “winner”—for both ideological and technological reasons. Ideologically, if everyone passed the tests there simply would be no way to justify elite status for any particular group: every student would qualify for the most elite colleges and jobs, thereby rendering the very hierarchy of elitism obsolete. A high-stakes, standardized test that everyone passed would then function to challenge White supremacy, not maintain it. Additionally, and technically speaking, in order to remain statistically valid standardized tests generally require that some portion of students fail the test (Popham, 2001).

These ideological and technical points are critical for understanding how high-stakes testing fits into the discourse about race and class issues in the current education reform movement in the United States, which relies explicitly on the rhetorical goal of closing racial and economic achievement gaps in test scores (Ladson-Billings, 2006). There is a great irony within this discourse of framing education reform around the closing of test-based achievement gaps. Because the ways standardized tests require some amount of failure, closing the achievement gap does not mean having everyone be successful on high-stakes, standardized tests. Rather, closing the achievement gap actually means
having proportional rates of failure and success amongst different groups. If education in the U.S. closed the high-stakes test score achievement gap amongst different groups it would simply mean that equal numbers of rich kids and poor kids pass and fail, equal numbers of White kids and African American kids pass and fail, etc. (Au & Gourd, 2013). Closing racial achievement gaps in high-stakes testing does not mean making sure all kids pass the tests, it simply means creating a racially proportionate underclass of students failing the tests.

**White Supremacist Impact of Modern Day High-Stakes Testing**

Up to this point in this chapter I have offered not only a history of high-stakes testing, but also an analysis of several key ways such testing operates in the maintenance of White supremacy. In the previous section I specifically highlighted how the concept of meritocracy has been coupled with the presumed objectivity of standardized tests to effectively mask structural racism and White supremacy within an explicitly race-neutral framework. In this section I turn to more empirical evidence of how high-stakes, standardized testing has a disparately negative impact on students of Color in classrooms and curriculum.

A look at the research evidence on the effects of modern day high-stakes, standardized testing on students of Color illustrates ways the tests function to support White supremacy in a variety of ways. Functionally high-stakes, standardized testing has served to support a White supremacist curriculum and classroom environment. Research has found that if a subject is not on the high-stakes test, it tends to get dropped from the curriculum that is taught in the classroom (Au, 2007). This phenomenon has a wide range of implications, but one of the critical issues that arises is that curriculum which explores issues associated with race and culture specifically is being ignored in response to the tests. Put differently, high-stakes testing evacuates multicultural, anti-racist perspectives out of the curriculum, creating a movement away from non-Whiteness towards a colorblind norm, a norm which ultimately supports Whiteness through racially disparate outcomes for students of Color (Au, 2009; Darder & Torres, 2004). Other research has found that high-stakes testing also promotes predominantly White, Eurocentric and Western views in some U.S. states’ world history, geography and U.S. history tests (Grant, 2001). Toussaint (2000-2001) tells the story of how, as an employee of a test scoring company, he was told to give higher scores to students who provided Eurocentric answers to test questions about Manifest Destiny and the colonization of the Americas. High-stakes, standardized tests also impose
Whiteness on teachers as well. Research has found that teachers who enter into the teaching profession with the stated goal of using multicultural, anti-racist content as a means to effectively teach the diverse students in their classrooms are giving up this goal in response to the curriculum pressures created by the tests (Agee, 2004). Since multicultural, anti-racist perspectives and content are not deemed legitimate by the tests, the end result is that within the high-stakes testing environment, multicultural, diverse, and non-White perspectives and content are being increasingly excluded from the classroom, thus acting as a tool for the maintenance of White supremacy.

The White supremacist curriculum enforced by high-stakes testing directly and negatively impacts students of Color. Research tells us that students learn best when they can connect themselves, their identities, their lives, and their experiences to their learning. This has proven to be true for students of Color in particular, especially those historically underserved by our school system: curriculum that connects to students’ cultures and identities fosters deeper connection to concepts and learning, and can lead to more academic success (Cabrera, Milem, & Marx, 2012; Romero, Arce, & Cammarota, 2009; Valenzuela, 1999). By legitimizing Whiteness through the delegitimization of non-Whiteness in curriculum and classroom environments, high-stakes tests explicitly include and exclude certain student identities in schools. Put differently, because high-stakes tests force schools to adopt a standardized, non-multicultural, curriculum that structurally enforce norms of Whiteness, it ultimately silences the cultures and voices of children of Color, particularly if those voices, cultures, and experiences are not contained on the tests. High-stakes testing thus requires non-Whiteness to be subtracted from the curriculum (McNeil, 2000; Valenzuela, 1999) because of its emphasis on the standardization of a norm that enforces Whiteness. Indeed, diversity itself has become a threat to survival and success within the systems of high-stakes testing because it is antithetical to the process of standardization and challenges the Whiteness communicated within what is considered the “norm.”

The White supremacy of high-stakes, standardized testing also manifests at the level of education policy implementation. At the state level the pressures of high-stakes, standardized testing are greatest in states with large, diverse populations of non-White students (Berliner & Nichols, 2005), and at the classroom and school level the narrowing of the curriculum is most drastic in schools with large, populations of non-White students. For instance, the Center for Education Policy (Rentner et al., 2006) found that 97 percent of high-poverty school districts, which are largely populated by diverse groups of non-White students, have instituted policies specifically aimed at increasing time spent on reading. This is compared to only 55% to 59% of wealthier,
Whiter districts. In this way, high-stakes testing has produced more restrictive, less rich educational environments for non-White students.

Further, high-stakes, standardized test scores are increasingly being used as part of the decision-making for school closures in the United States, and these closures have disproportionately impacted communities of Color. For instance, according to the National Opportunity to Learn Campaign (2013), the population of the 54 public schools to be closed in Chicago in 2013 were 88% African American and 94% low income. In New York City, the population of the 26 public schools to be closed was 97% African American and Latino, and 82% low income, and in Philadelphia the 23 schools to be closed were 96% African American and Latino and 93% low income. School closings, as part of an extension of the use of standardized test scores to make high-stakes educational decisions, clearly function in a way that support institutionalized racism against working class Black and Brown kids.

**High-Stakes Testing, Discipline and the Schools-To-Prisons Pipeline**

High-stakes, standardized testing also works within a system of education policy that seeks to discipline those that do not “fit” within the norms, and this too also works in the interest of White supremacy. High-stakes testing establishes criteria through which those with power can define “failure” (transgression from the norm) and “success” (conformity to the norm), and then use those designations to rank students, teachers, schools, states, and countries. This enables the surveillance of those who do not “fit” within certain test-defined boundaries of success. This ever increasing web of surveillance (Hanson, 2000) works, in part, because categorically defined deficiencies, according to Lipman (2004), are made, “visible, individual, easily measured, and highly stigmatized within a hierarchical system of authority and supervision” (p. 176).

The use of high-stakes, standardized tests as a tool to discipline through the evaluation and ranking of students, teachers, schools, and administrators enforces White supremacy in a very concrete way as well: it directly contributes to the schools-to-prisons pipeline. This happens along two pathways. The first is in the kind of school culture fostered by using high-stakes, standardized tests to evaluate teachers and students. Functionally such testing promotes a school culture of discipline and punishment (Foucault, 1995). Students and teachers are in essence surveilled by the tests, and when they step out of line by not performing well, they are punished. In this sense, within a high-stakes testing environment, teaching and learning are done
High-Stakes Testing

under constant threat and fear. Not only is threat and fear not conducive to learning, such conditions also acculturate students to accepting more prison-like school culture. Further, this overall view of test-defined success and failure functionally pathologizes those who fail, and as I have discussed at length here, working class communities of Color have had persistently lower test scores historically. Thus, as Ekoh (2012) explains, this pathologizing of students and failure relates directly to the systematic functioning of institutions:

If unable to thrive within the confines of the system, they [students] are pathologized and condemned as incompetent and lazy, for its not the system that fails them, but they, that fail the system. As such they must be re-made “healthy” and “reintegrated” into the oppressive status quo. Healthy, in this context referring to the state of mind that results when one has been successfully disciplined and “indoctrinated” into the dominant and oppressive system in such a way as to lose one’s critical consciousness…(pp. 70–71)

Ekoh’s point is critical for understanding the racial and racist relations I am attempting to pinpoint in my discussion here. The pathologizing and condemning of some groups within a system (in this case, education and schooling) simultaneously requires that other groups be deemed healthy and acceptable within that same system. Thus, if low income students of Color are pathologized, it is not simply a matter that White students are left to navigate a level or neutral educational playing field. Instead I’m suggesting that, in their function of supporting White supremacy, high-stakes standardized test also provide a pathway to success for affluence and Whiteness at the cost of the failure of low income students of Color. The system of testing is fundamentally structured not to support test-defined “success” for all, and instead requires the failure of some to prove the worthiness of others. The tests’ role in this type of differentiation, or “distinction” in Bourdieu’s (1984) terms, is illustrated by research that has found that high-stakes, standardized testing can contribute directly to the actual imprisonment of students of Color: a study by the National Bureau of Economic Research (Baker & Lang, 2013) found that standards-based high school exit exams increase the rate of incarceration by 12.5%. Put differently, kids who fail the exit exams (who are disproportionately non-White) have their options in life greatly limited, and one of the end results is a 12.5% increase in the chances that they might end up as a part of the criminal justice system. As this study illustrates, students who fail tests have opportunity taken away while passing students are allowed more opportunity. Given the racist and White supremacist functioning of the tests, this allows us to see how the tests are not neutral, and instead are an integral part of a system that supports affluent, White success.
**Who’s Benefitting from Testing?**

A critical question to ask of all policies is, “Who benefits?” (Brantlinger, 2006). As I have argued in this chapter, the short answer is “Whiteness” and by extension, White people, if for no other reason than the role that high-stakes, standardized testing selectively supports the upward mobility of affluent White folks over others. However, there is another, very concrete answer to the question of, “Who benefits from the implementation of high-stakes testing?” According to the National Center for Educational Statistics (2013), the total expenditures for all education in the United States (public and private, from prekindergarten through graduate school) were estimated to be $1.2 trillion for the 2011–2012 school year, with U.S. elementary and secondary school expenditures being $700 billion of that total.

Free-market loving, corporate education reformers are very excited about the possibility of tapping into this huge market through the delivery of tests aligned to new standards, amongst other profit-making opportunities (Ravitch, 2013). This profit making orientation is illustrated by Joanne Weis (2011) in Ravitch (2013), who formerly served as the director for the Obama Administration’s Race to the Top initiative and is currently U.S. Secretary of Education Arne Duncan’s chief of staff, when she stated:

> The development of common standards and assessments radically alters the market for innovation in curriculum and development, professional development, and formative assessments. Previously these markets operated on a state-by-state basis, and often on a district-by-district basis. But the adoption of common standards and shared assessments means that educational entrepreneurs will enjoy national markets where the best products can be taken to scale. (p. 17)

High-stakes, standardized testing is one part of the free-market education reform puzzle, and there is money to be made. In his analysis for the Brookings Institute, Chingos (2012) offers an estimate of $1.7 billion being spent nationwide on the tests alone. This number is conservative, however, because it focuses on just the cost of the tests themselves and does not account for test-aligned textbooks, technology, and professional development often devoted to specific assessments. As Cavanagh (2013) explains, one estimate for technology-based testing and off-shoot technology requirements to give the tests or prepare for the tests (e.g., new computers or tablets) was $1.6 billion for the year 2011, and that number was estimated to increase by 20% for 2012. According to Cavanagh, another estimate suggests that the total market for K-12 testing in 2012 was $3.9 billion, and it was expected to increase 4–5% every year to $4.5 billion by 2014. Within this mix companies
like Pearson and McGraw-Hill are profiting both in developing and selling tests themselves (including administrative services like grading, analyzing, and reporting information) but also in selling textbooks and providing other services and materials in support of test preparation (Figueroa, 2013). So when we post the question of, “Who benefits from high-stakes, standardized testing?” in addition to the maintenance of status and upward mobility of affluent Whites, we also see the testing industry having enormous benefit for educational corporations in the U.S.

**Resistance to High-Stakes, Standardized Testing**

Despite the ubiquity and hegemony of high-stakes, standardized testing in education policy today, it is important to understand that there has always been significant resistance. This resistance, however, has taken different forms in different times, and it was early resistance that looked to challenge the White supremacy of the tests most explicitly. For instance, Mexican American and African American educators critiqued the standardized test-based eugenics and I.Q. testing movements in the U.S. One of the earliest African American educators to publicly challenge the findings of prominent psychologists involved in the I.Q. testing and eugenics movements was Horace Mann Bond, the Director of the School of Education at Langston University in Oklahoma. In 1924 Bond critiqued I.Q. testing and eugenics in *Crisis*, the magazine of the National Association for the Advancement of Colored People (Au, 2009). George Sanchez (1940) argued about the racial bias in testing relative to Mexican American students in New Mexico, and in 1940, W.E.B. Du Bois, as quoted in Guthrie (1998), also recalled:

> It was not until I was long out of school and indeed after the (first) World War that there came the hurried use of the new technique of psychological tests, which were quickly adjusted so as to put black folk absolutely beyond the possibility of civilization (p. 55).

However, this early resistance to I.Q. tests by educational leaders of Color did not prove to be enough to stop the ascendancy of standardized tests as a tool of White supremacy in the United States.

More recently there has been a surge of resistance to high-stakes, standardized testing, but this resistance has been more general in nature and less focused on the racism and White supremacy associated with the tests. One such organized resistance to the White supremacy of high stakes testing took place in 2003, when two community activist groups in California, Justice Matters and Youth Together, organized a campaign for the establishment of
an alternative to the California High School Exit Exam (CAHSEE) for students in the Richmond School District, specifically because of the disparately racist impact of the CAHSEE on youth of Color and immigrant youth there. While the immediate campaign was unsuccessful in terms of changing policy, it was very successful at gathering students, parents, and community groups together as a coalition pushing for racial justice in California State testing (Gray-Garcia, 2012).

There have also been several individuals resisting the tests in the last decade, but this resistance has not been racially based and instead has focused on how high-stakes testing corrupts best practices for teaching and learning. In 2007, middle school teacher David Wasserman personally boycotted the Wisconsin Knowledge and Concepts Exam, but eventually gave in to district threats to his job and administered the test. In 2008 Seattle Public Schools middle school teacher Carl Chew also refused to give the Washington Assessment of Student Learning (WASL), and the district suspended him for his resistance. Juanita Doyon, a parent in Washington, also started a group, Mothers Against WASL, to organize Washington parents against what she saw as an assessment that was detrimental to education and learning. North Carolina teacher Doug Ward similarly boycotted the state test there on the grounds that it was inappropriate for students with disabilities, and he was fired for his actions (Oulahan, 2012). These smaller and individual acts of resistance, while not all “successful”—especially given the punishments resisters faced—were critical because the local and national stir created by their acts proved to be important for developing the community conversation around the broader issues of testing and educational justice.

In the last few years, however, particularly since the election of President Obama, the establishment of the Race to the Top initiative of his administration, and the implementation of the Common Core State Standards, resistance to high-stakes, standardized testing has grown exponentially. Perhaps most famously, in January of 2013 all of the teachers at Seattle’s Garfield High School boycotted the MAP test, causing a national wave of support both for their particular movement, but also around the problems with high-stakes testing generally. Teachers at six other Seattle public schools joined their colleagues at Garfield, and many saw this boycott as the beginning of the next wave of the movement against testing (Hagopian, 2013). While some of the leadership for the MAP test boycott clearly connected this test to the issues of racism and White supremacy (Hagopian, 2011), teachers boycotted this test for a variety of reasons not connected to race at all (Hagopian, 2013).

Indeed, 2013 saw a significant uptick in testing resistance generally. As FairTest (2013) summarizes in their report of test resistance at the time:
Students, parents and teachers engaged in boycotts, “opt-out” campaigns, and walkouts in Seattle, Portland, OR, Denver, Chicago and New York, with smaller events in other communities across the nation. Seattle teachers in several schools boycotted one test; most students, backed by their parents, did not take it when administrators gave it. Families boycotted at 40 schools in New York City and many outside the City…

Providence students staged a “zombie march,” then organized a group of prominent adults to take the state graduation test, a majority of whom failed the math. Another effective action was a “play in” at Chicago Public Schools offices protesting the dozen or more standardized tests given to Kindergarten, first and second grade students. Demonstrations and rallies in New York, New Jersey, Texas, Chicago and elsewhere also garnered significant attention…

Legislative efforts took center stage in Texas, where parents led a successful charge to reduce high school end-of-course graduation tests from 15 to 5. In Minnesota, teachers unions and civil rights groups backed legislation that repealed the state’s graduation exams. Bills were introduced in other states as a first step. Other concrete gains included a modest roll-back of testing in the early grades and high school in Chicago and of some Seattle tests. (n.p)

Facebook groups of teachers fed-up with testing and other aspects of the current education reform movement are growing, one of the biggest being the Badass Teachers Association (2014) which boast upwards of 38,000 supporters. Further, coalition groups of parents, teachers and professors, like United Opt Out National (2014), are pushing to have their children “opt out” of taking the tests, much to the chagrin of administrators, officials and policymakers. In February of 2014, for instance, an alliance of Chicago parents started a boycott of the ISAT exam, as announced by another test resistance group, More Than a Score. This boycott launched with 500 students across 29 Chicago schools (Reid, 2014), and teachers at two Chicago public schools, Drummond Montessori and Saucedo Academy, voted as an entire faculty to boycott the same tests, despite promised reprisals from the district (Ahmed-Ullah, 2014). A collection of organizations nationally, including FairTest, Network for Public Education, Parents Across America, Save Our Schools, and United Opt Out National, along with several prominent anti-testing individuals, also launched a campaign and network called the “Testing Resistance & Reform Spring: Less Testing, More Learning,” to continue to generate resistance to the testing in a more unified and coordinated way (Testing Resistance & Reform Spring, 2014). Thus, there is a significant test resistance movement growing (Rethinking Schools, 2014), even if much of that movement is not necessarily based on the pushback against the racism and White supremacy that high-stakes, standardized test enforce within our system of education in the United States.
Conclusion

High-stakes, standardized testing has become universal in education policy, reform, and practice, and this universality has been established largely on commonsense understandings that avoid the faulty logics that lay at their foundations, deny the White supremacist history of such tests and gloss the racist outcomes of testing today. In this way we could argue that the imposition of high-stakes, standardized testing on public schools in the United States mirrors our larger discussion about education reform more generally: despite all prevailing evidence about the ineffectiveness of these reforms, including ample evidence that these reforms are hurting kids and communities of Color, these backward policies continue to be advanced by politicians, pundits, and corporate leaders. Meanwhile the very real and very racist structural social, cultural, and economic inequalities that we know have a disproportionately high impact on the successes and failures of students of Color go untouched within mainstream public and political conversations.

However, as the testing juggernaut speeds ahead, one thing should give us all hope. The failures of the No Child Left Behind era are clear, the price of conceding to the requirements of Race To The Top is being questioned, and the popular pushback against the Common Core Standards (along with the impending threat of national testing that those standards bring) have all come together to shift the commonsense consciousness of many parents, teachers and students. Locally and nationally, parents, teachers and students—and yes, communities of Color included—are calling out injustices created by these tests (Rethinking Schools, 2014), and although inconsistent, these calls increasingly include the structural racism and White supremacy that high-stakes, standardized testing fundamentally supports.

References


Gray-Garcia, T. (2012). Teaching is not testing: A community-led struggle to find an alternative to California’s graduation exam. In W. Au & M. Bollow Tempel (Eds.), *Pencils down: Rethinking high-stakes testing and accountability in public schools* (pp. 183–188). Milwaukee, WI: Rethinking Schools, Ltd.


Oulahan, A. (2012). None of the above: Defiant teachers show they have had enough of high-stakes testing. In W. Au & M. Bollow Tempel (Eds.), Pencils down: Rethinking high-stakes testing and accountability in public schools (pp. 189–193). Milwaukee, WI: Rethinking Schools, Ltd.


**Note**

1. I would like to be clear on how I am using the term “White supremacy” and how I am using racial categories like “White” in this chapter. The kind of White supremacy I refer to here speaks to what I see as the White supremacist outcomes of systematic and institutional racism in the United States in particular. As I argue here, as the expression of education policy and a pivotal apparatus for leveraging education reform in the United States, high-stakes, standardized testing functions as a tool for the maintenance of White supremacy. Additionally I want to be explicit in that I use “race” here not as a biological construct—since it has no factual basis in biology (Gould, 1996)—but as a social, historical, and political construction (Omi & Winant, 1994) that is used to sort and define human populations based on their perceived, and sometimes legally defined, racial identity in the United States.