Critical Education

Volume 1 Number 10 December 20, 2010 ISSN 1920-4125

How Well Does the Standards Movement Measure Up?

An Analysis of Achievement Trends and Student Learning, Changes in Curriculum and School Culture, and the Impact of No Child Left Behind

Lawrence C. Stedman
State University of New York at Binghamton

Citation: Stedman, L. C. (2010). How well does the standards movement measure up? An analysis of achievement trends and student learning, changes in curriculum and school culture, and the impact of No Child Left Behind. *Critical Education, 1*(10). Retrieved [date] from http://m1.cust.educ.ubc.ca/journal/index.php/criticaled/issue/view133

Abstract

This is the first of two papers examining the standards movement. In it, I review data from NAEP, the SAT, the international assessments, transcript studies, and NCLB assessments, as well as surveys and case studies of changes in curriculum and pedagogy. The picture is a bleak one. Over the past quarter century, achievement has stagnated, dropouts and aliteracy have grown, and large minority achievement gaps have persisted. The quality of student learning remains poor. School changes, stratified by class and race, have constricted instruction and harmed students and teachers. NCLB has made things worse, not better. Even in the two areas where the movement has achieved some success—math achievement at the lower grades and high school academic enrollments—the gains were largely superficial, other forces such as teaching-to-the-test and social promotion contributed, and serious deficiencies remain.

In the second paper, "Why the Standards Movement Failed," I examine the educational and political reasons for the failure—including its misconstruction of pedagogy and links to the neoliberal reform project—and propose a progressive alternative.

Introduction

For a generation, the United States has pursued an aggressive program to raise standards and improve achievement. The 1983 report, *A Nation at Risk*, galvanized the movement. To address the "rising tide of mediocrity that threatens our very future as a Nation and a people," it proposed "rigorous and measurable standards," stronger graduation requirements, and standardized tests "at major transition points" (National Commission on Excellence in Education, 1983, pp. 4, 70, 73). States responded by expanding testing programs and requiring more academic coursework for graduation (Lillard & DeCicca, 2001, pp. 460–461; Snyder & Hoffman, 1991, pp. 139–143). During the 1980s, districts across the U.S. adopted the effective schools formula, emphasizing discipline, high expectations, and testing (Stedman, 1987). To spur further reform, Congress expanded the National Assessment of Educational Progress (NAEP) to state testing and reporting and the National Council of Teachers of Mathematics issued its groundbreaking *Curriculum and Evaluation Standards for School Mathematics* (NCTM, 1989).

During the 1990s, other professional organizations, funded by the federal government, issued standards documents in the major subjects (Kendall & Marzano, 2004). Although designed to improve curricula and pedagogy, they soon became entangled with testing and accountability reforms. The Clinton administration's Improving America's Schools Act of 1994 (IASA) required states to develop standards, test students annually, and report data by race and ethnicity. It was this act, not NCLB, that introduced the *adequate yearly progress* benchmark (Dillon, 2009; IASA, 1994). In 1994, Congress also established the Goals 2000 program, which called for U.S. students to be "first in the world in mathematics and science achievement" (National Education Goals Panel, 1999, p. vi). In its annual reports, the NEGP monitored achievement trends and reported on the test-score gaps between white and minority students.

The capstone of these efforts was the widely supported, bipartisan No Child Left Behind Act of 2001 (NCLB, 2002). Stressing "stronger accountability for results" (U.S. Department of Education, n.d.), NCLB requires states that accept Title 1 funds to test reading and math annually in grades 3–8 and once in 10–12, and science at three school levels. States must set performance standards for each grade and ensure that *all* students are proficient within 12 years. Schools are required to make adequate yearly progress overall and with minority, limited English proficiency, and low-income students. The consequences for failure include mandatory improvement plans, transfer rights, and school takeovers.

The standards movement, therefore, has been intense and widespread, and has shaped education for the past 25 years. What has happened? How much has student achievement improved? What progress has been made in closing the minority achievement gap? How have the lives of students and teachers been affected? What are the forces behind these effects and what are their implications for school reform?

I address these questions in two papers. This one focuses on the empirical evidence and is divided into two parts. Part I covers the test-score evidence, including NAEP, the SAT, and the international assessments. Part II focuses on changes in curriculum and school culture, using findings from case studies, transcript studies, and national surveys. In the second paper, I shift to an educational and political diagnosis of the standards movement itself and what has gone wrong. In it, I discuss the movement's transformation into a bureaucratic accountability system and its fundamentally flawed approach to schooling. I examine its relationship to the neoliberal reform project and construct an alternative approach to reform.

Part I: In Search of Excellence

Achievement Trends

NAEP

NAEP provides our best measure of national achievement trends (Stedman, 2009). It regularly tests large, nationally representative samples of students. It uses authentic materials, such as excerpts from literature and historical documents, and asks students to make inferences and apply knowledge. Since 1969, it has tested over 20 areas. In the 1980s, NAEP was split into two—a long-term trend program to continue tracking 9, 13-, and 17-year-olds in reading, math, and science and a main program to assess 4th, 8th, and 12th graders in these and other areas over shorter periods. (For a time, the trend program also tracked writing by grade level.)

The long-term patterns are striking. Since the early 1980s, achievement has generally stagnated (see Figure 1). Although there have been some math gains at the lower grades (their quality is a matter I will return to later), high school performance is a better gauge of the overall impact of K–12 education, and it remains largely unchanged.

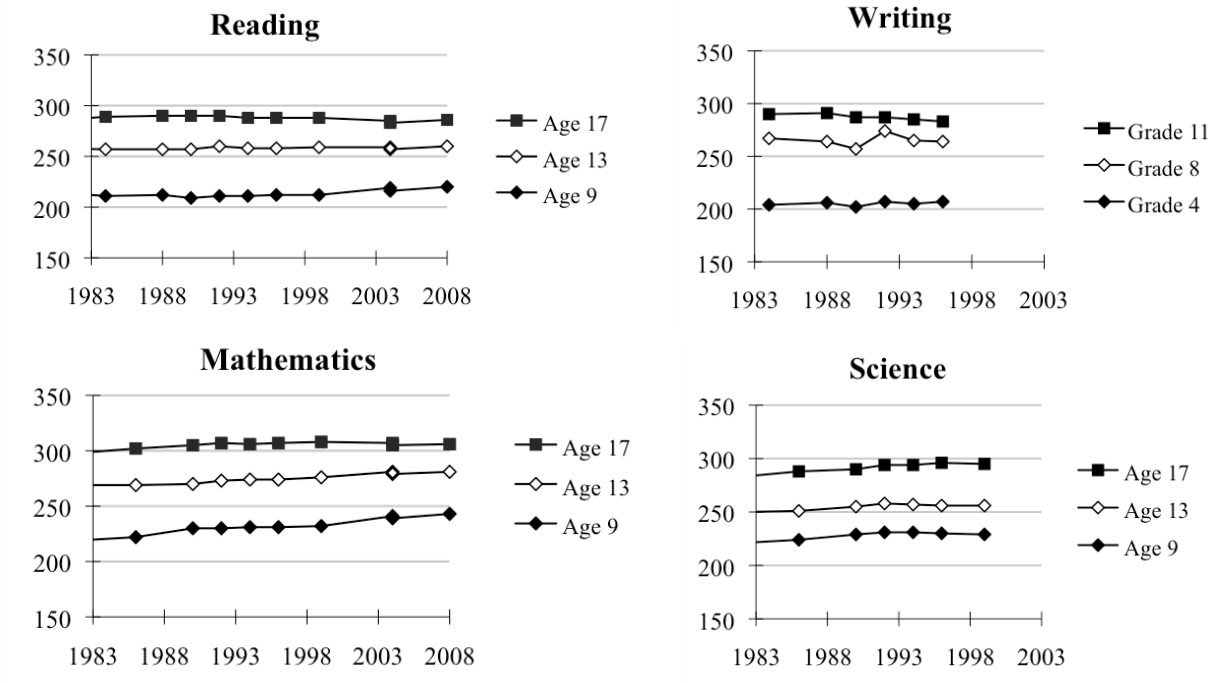

Figure 1. Trends on the long-term NAEP assessment. Science and writing were dropped after 1999. (Writing was not reported in 1999 due to technical problems.) The NAEP scale is 0–500, yet 150–350 is a fairer representation. Two scores are plotted for 2004; from the old and revised assessments. (The differences in national averages were only 2-3 points.) Data are from Rampey, Dion, & Donahue, 2009, pp. 9, 29 (reading, math); Campbell, Hombo, & Mazzeo, 2000, p. 9 (science); and Campbell, Voelkl, & Donahue, 1998, p. vi (writing).

Reading and writing trends are illustrative. Literacy is the cornerstone of learning, but for the past two decades, middle school and high school reading have languished in both assessments (Grigg, Donahue, & Dion, 2007, pp. 4–5; National Center for Education Statistics, 2010b, p. 25). There were even some small declines among high school seniors. Writing did not improve at any level of schooling in the 1980s and 1990s and, since then, on the main NAEP, gains and fluctuations have been minor ones (Salahu-Din, Persky, & Miller, 2008, pp. 8–9, 36–

37). Only about a third of 12th graders are considered proficient readers, and less than a fourth, proficient writers (Grigg et al., 2007, p. 5; Salahu-Din et al., 2008, p. 37).

Goals 2000 and NCLB set science achievement as a vital national goal yet here, too, the news is not good. After some gains in the 1980s, science performance stalled in the 1990s (see Figure 1). Since then, in the main NAEP, 8th graders have made no progress, while 12th graders have lost ground and remain below where they were in 1996 (Grigg, Lauko, & Brockway, 2006, pp. 19, 30–31). For seniors, NAEP also reported: "Declines seen in all fields of science" (p. 31).

The standards movement has fared little better in geography or U.S. history (see Figure 2). Civics achievement, which has been assessed for 40 years, has also stagnated; there have even been some declines along the way (Lutkus & Weiss, 2007, p. 7; Stedman, 2009, pp. 13–14).

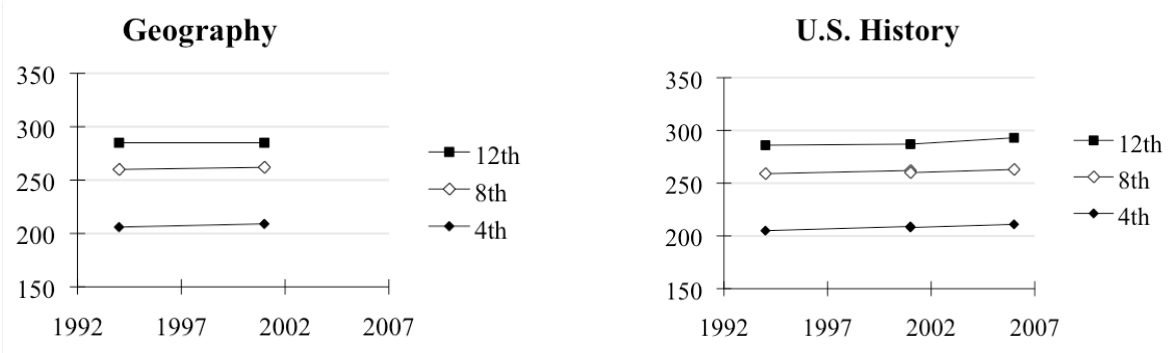

Figure 2. Trends on the main NAEP assessment. Geography was last assessed in 2001. Geography data are from Weiss, Lutkus, Hildebrant, & Johnson, 2002, p. 18. U.S. history data are from Lee & Weiss, 2007, pp. 7–9. NAEP now uses testing accommodations. Two scores are plotted for U.S. history in 2001, one with accommodations, one without. This made little difference. The national averages were identical at 12th; 1-2 points different at 4th and 8th.

SAT Scores

Since *A Nation at Risk,* the math SAT gain has been small, similar to what the typical test-prep program produces (Camara & Powers, 1999), while verbal scores have not improved at all (see Figure 3). The trends are lackluster, especially given the changes in the test-taking pool. As in the past, it grew in size and diversity but, this time, it became more advantaged. Over the past 27 years, the family income and educational background of those taking the SAT have improved substantially (College Board, 2010, pp. 3–4; Ramist & Arbeiter, 1983, pp. 3–4).

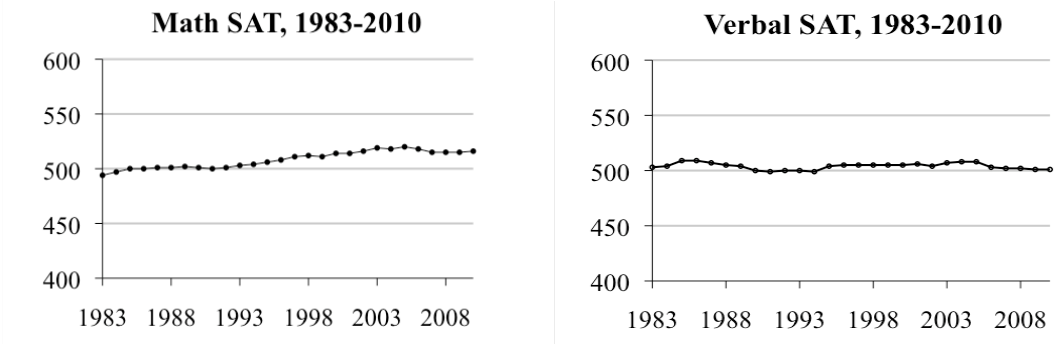

Figure 3. Data are from College Board, 2010. All scores are on the re-centered scale. The College Board converted pre-1996 scores (and those of a few individuals who took the SAT from 1996 to 1999) to the new scale.

In any event, the SAT is a poor barometer of school performance (Stedman, 1998). It judges a self-selected group of *academically elite* students (most report ranking in the top fifth of

their high school classes and having taken 4 or more years of each academic subject) (College Board, 2010, pp. 5–8). It also measures little of the high school curriculum. As Lynne Cheney (n.d., p. 8) once aptly put it,

> Looming over our educational landscape is an examination that, in its verbal component, carefully avoids assessing substantive knowledge gained from course work. Whether test-takers have studied the Civil War, learned about Magna Carta, or read Macbeth are matters to which the SAT is studiously indifferent.

Even the College Board has, for years, warned against using it to measure school quality, labeling it "not valid" for such purposes (College Board, 1996, p. iii; 2000, p. iii; 2010, p. 2).

International Assessments

During the past decade, three major international assessment programs have reported achievement trends involving U.S. students (see Appendix for details). While such assessments have their limitations, they are the result of decades of experience and evaluators have developed well-honed methods to control potential problems (Gonzales et al., 2008). Many popular criticisms of the assessments—such as unfair mass vs. elite comparisons—no longer apply. Others, such as test and curricular biases against the U.S., have not held up (Stedman, 1994, 1997). Even in the past, researchers reached credible judgments by focusing on the top 10% or on similarly small groups of advanced, college-preparatory students—or on the test items that U.S. teachers reported covering.

So, what have the recent trends been? Since 1995, U.S. 4th and 8th graders gained modestly in math, but less in science, where 4th graders even slipped (Gonzales et al., 2008, pp. 8, 33). 8th graders were not tested in reading, but 4th graders' reading scores have been declining since 1991 (Martin, Mullis, Gonzalez, & Kennedy, 2003, p. 15; Mullis et al., 2007, p. 44). In each area, NCLB's impact can be questioned. 8th graders made most of their math gain over a decade ago (1995–1999), *before* NCLB was passed (cf. Mullis, et al., 2004, p. 42 and Gonzales et al., 2008, p. 8). Their science scores declined *after* it was enacted (cf. Martin et al., 2004, p. 44 and Gonzales et al., 2008, p. 33), as did 4th graders' reading scores.

U.S. high school students also lost ground at different points. From 1991 through the early 2000s, their reading performance worsened (Elley, 1992, p. 24; OECD, 2010, p. 39), while in the first half of the 2000s, their math and science scores declined substantially (Baldi et al., 2007, pp. 47–48). From 2006 to 2009, there were sudden jumps in both subjects (OECD, 2010, pp. 156, 159), but countries often have erratic changes in their scores and such rapid upswings suggest curriculum drilling or sampling fluctuations rather than genuine improvement. Although Goals 2000 called for U.S. students to be #1 in math and science, they still lag well behind those in many countries (Baldi et al., 2007, pp. 6, 12; OECD, 2010, pp. 156, 159). The concern is not the lack of U.S. supremacy. The point is that, during much of the standards era, U.S. achievement *declined* rather than improved internationally in key areas, especially literacy, and that, in spite of an intense test-driven intervention in schooling, U.S. students remain seriously deficient in skills and knowledge (Baldi et al., 2007; Lemke et al., 2004, pp. 7–12; Stedman, 2009). A historical perspective is illuminating. The weaknesses today are as deep and troubling as those in the international assessments of the 1980s and 1990s (Stedman, 1994, 1997). The last time 12th graders were assessed, in 1995, even some of the most-academically talented U.S. students, those taking physics and advanced mathematics, were doing poorly (see Appendix).

The massive standards push after *A Nation at Risk*, therefore, has made little difference in achievement, especially at the high school level. Even so, the U.S. Department of Education and major organizations proclaimed the success of school reform. Secretary of Education Spellings heralded NAEP trends as "proof that *No Child Left Behind i*s working" (U.S. Department of Education, 2005) and "that we are on the right track" (Glod, 2009). The National Council of Teachers of Mathematics (2004, 2005) touted "steady improvement" in math scores and attributed it to the standards it had introduced in 1989. I examine their claims in later sections.

Tinkering Towards Equality

Crediting NCLB, President Bush and Secretary Spellings also trumpeted a narrowing of the minority achievement gap (Bush, 2008; U.S. Department of Education, 2005). They focused, however, on recent changes among the youngest students and ignored the overall picture (see Appendix). In a diverse array of subjects, the gaps remain vast and, especially among older students, are similar to those of the past. In the long-term NAEP, for example, most gaps today are as large as they were in the late 1980s (1986–1990) (Rampey et al., 2009, pp. 15, 17, 35, 37). In the main NAEP, most are about as large as when trend lines began (1990–1994) (Grigg et al., 2007, p. 6; Lee & Weiss, 2007, p. 14; NCES, 2009, pp. 10, 25; NCES, 2010c, pp. 10, 27; Weiss et al., 2002, p. 28). Science gaps generally have been the largest and have not closed perceptibly in several decades (Campbell et al., 2000, pp. 39–40; Grigg et al., 2006, pp. 20, 32).

Even among the youngest students, the picture is not so rosy. Despite gains, black and Latino students remain about as far behind white students in mathematics as they were in the mid-1980s (Rampey et al., 2009, p. 34). In reading, while gaps have closed somewhat, black and Latino 9-year-olds still have not caught up with white 9-year-olds of *forty years ago* (Rampey et al., 2009, pp. 14, 16). Furthermore, since NCLB was enacted, progress in both areas has slowed greatly (NCES, 2009, pp. 10, 25; NCES, 2010c, pp. 10, 27; Rampey et al., 2009, pp. 14, 34)

The high school situation is dismaying. In the latest NAEP assessments, *four-year* lags are the norm (see Figure 4). For example, *black and Latino 17-year-olds barely match the scores of white 13-year-olds*. In fact, in more than a generation of testing, they have seldom performed much better than white 13-year-olds in *any* subject (Stedman, 1998).

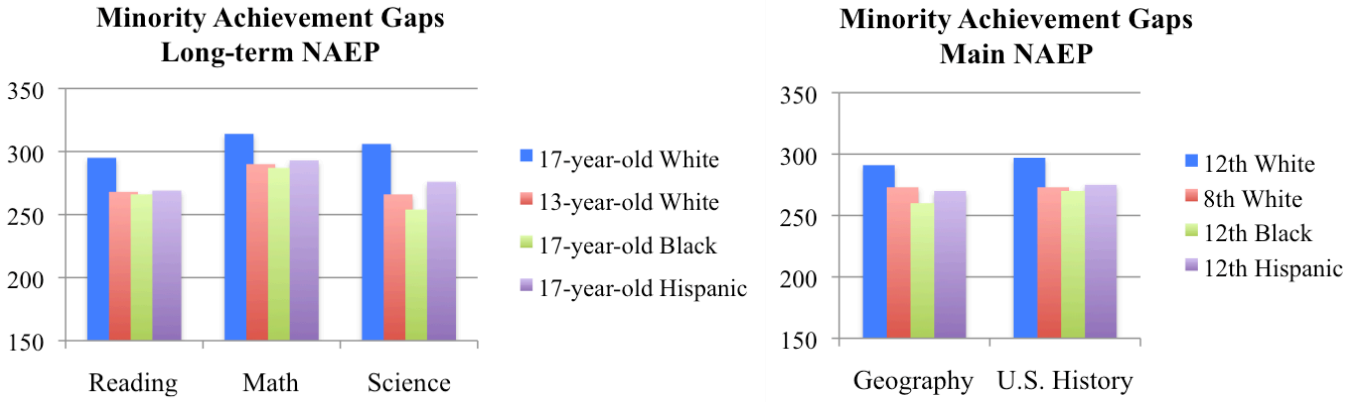

Figure 4. Gaps by race and ethnicity from the most recent assessments with comparable data. Reading and math data are from 2008; science, 1999; geography, 2001; and U.S. history, 2006. NAEP uses the classification "Hispanic" not Latino. Data are from Rampey et al., 2009, pp. 15, 17, 35, 37 (reading and math); Campbell et al., 2000, p. 37 (science); Weiss et al., 2002, p. 27 (geography); and Lee & Weiss, 2007, p. 13 (U.S. history).

At one point in the past, black and Latino high school students had made substantial strides in reading and math, but the timing of their gains is surprising (see Figure 5).

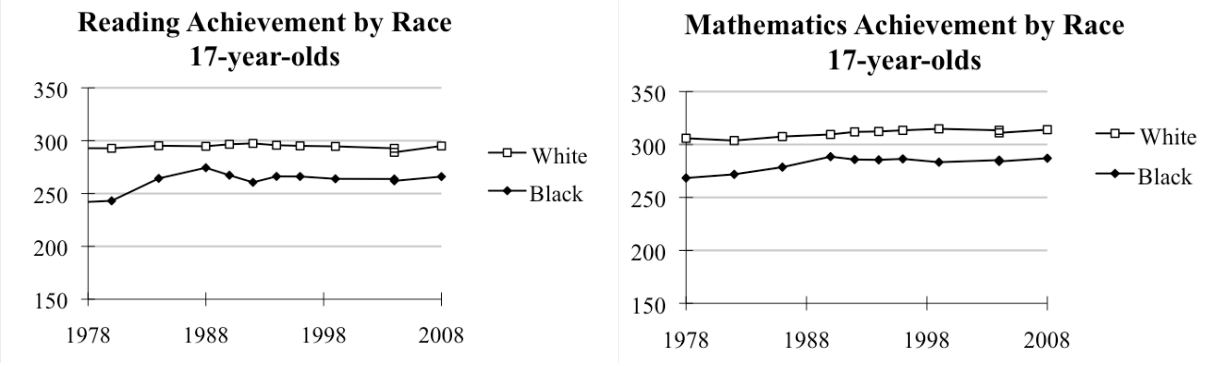

Figure 5. Trends on the long-term NAEP assessment. Two scores are plotted for 2004, from the old and revised assessments. Data are from Rampey et al., 2009, pp. 15, 35. For more long-term graphs, see Stedman, 2009.

Their scores were improving *before A Nation at Risk*. The gains continued during the Reagan years, but progress in closing the gaps abruptly ended with the Clinton-Bush era. Crediting the standards movement for this ancient improvement would be a stretch. It was more likely created by the anti-poverty, civil rights, and educational programs of an earlier era (Orfield, 2006, pp. 6–7). In any case, whatever impact the standards movement initially had on the equality of high school achievement, its effects have largely petered out in recent decades.

NCLB: Full of Fury

Despite grand claims about NCLB's success, many evaluators have determined it has not improved achievement and may have dampened progress (Fuller et al., 2007; Lee, 2006). NAEP data show this well. Since 2002, the nation's math gains in the early grades have slowed greatly, while those in high school have stagnated (NCES, 2009, pp. 8, 23; Rampey et al., 2009, p. 29). Reading scores have changed marginally—in the main NAEP, they are virtually frozen in place (Grigg et al., 2007, p. 5; NCES, 2010c, p. 1; Rampey et al., 2009, p. 9). With few exceptions, minority achievement gaps have closed little or not at all as the result of NCLB (see Appendix). In the long-term NAEP, for example, the changes in the reading and math gaps between 2004 and 2008 were tiny at *all* ages (some gaps even increased) (Rampey et al., 2009, pp. 14–17, 34–37). *State* testing programs have shown some gains, but those are often short-term and inflated compared to NAEP's (Hoff, 2007).

Veteran policy analysts also have questioned NCLB's effectiveness. After examining state and national NAEP data, Smith (2007) concluded that

> These findings not only contradict Secretary Spellings' contention that NCLB is working, but they also suggest that all the 1999–2004 gains that she touted came in the first three years of the period, before NCLB was enacted. (p. 4)

He also summarized 13 studies showing "evidence of effectiveness largely lacking" (p. 4) and sharply questioned the congressional rush to reauthorize it. Orfield (2006), the head of Harvard's Civil Rights Project, reached similar conclusions. He observed that

> the current trends will leave the nation very far from reaching the 100%

> proficiency goal. In Shakespearean terms, we've been experiencing a massive process "full of sound and fury, signifying nothing." (p. 6)

I assess NCLB's impact on student proficiency and school climate later in the paper.

Reflections on the Findings

The Standards and Mathematics

Surely, standards advocates will feel I have given short shrift to the math gains at the lower grades. It is hard, though, to consider a movement successful, let alone on the right track, when only one subject has markedly improved—and only at younger ages—and most other subjects and grades remain in the doldrums. Even the math gains turn out to be less impressive when scrutinized. In the main NAEP, 4th grade math scores have gone up 27 points, while 8th grade scores have risen 20 (NCES, 2009, pp. 8, 23). But this is on a 500-point scale and has taken almost 20 years to accomplish (1990–2009). The trends are gradual but, in NAEP graphs, they look like dramatic increases because the scales are improperly truncated—showing only 30 points of the 500-point scale (NCES, 2009, pp. 8, 23). Furthermore, the gains in the long-term NAEP have been only half as much during this time (Rampey et al., 2009, p. 29).[1]

At first glance, the younger students also seem to have made a stunning improvement in math proficiency. From 1990 to 2009, in the main NAEP, the percentage of 4th graders achieving proficiency in math tripled to 39%, while that of 8th graders more than doubled, from 15% to 34% (NCES, 2009, pp. 8, 23). The annual growth rates, however, are quite modest: only 1 to $1\frac{1}{3}$ percentage points. At those rates, it will take several generations, from 45 to nearly 70 years, to bring all students to proficiency, the stated objective of Goals 2000 and NCLB. Given the recent slow downs in their math scores, it will likely take even longer.

The connections between new or upgraded curricular standards and the math gains are not clear-cut. In the decade after NCTM issued its standards, younger students gained little on the long-term NAEP—only 2 points at age 9 and 6 points at age 13 on the 500-point scale (Rampey et al., 2009, p. 29). While their scores rose more on the main NAEP, the gains in both assessments were partly caused by demographic changes in test-takers rather than real improvement in math learning. Over time, as education levels in the society have risen, NAEP test-takers have come from better-educated families, helping drive scores up (Braswell et al., 2001, p. 246; Rampey et al., 2009, p. 54). The scores also have been artificially inflated by test drilling. It was a logical response to NCLB and the pressures of increased state testing in the early grades (Orfield, 2006). One sign of it was that the younger students' math scores jumped suddenly between 1999 and 2004, but have improved only modestly since then (NCES, 2009, p. 8; Rampey et al., 2009, p. 29).

Notwithstanding their gains, younger students continue to do poorly in math. Most 8th

---

[1] The discrepant results do not seem related to the assessments' test content or curricular alignment. Panels of math educators and teachers have judged the main and long-term NAEP items equally relevant to the curriculum and to have received similar classroom coverage (Pellegrino et al., 1999, pp. 77-78). Like the main NAEP, the long-term NAEP has evolved over time, with updated frameworks, items, and materials (Stedman, 2009).

graders, including *many* of those deemed "proficient," still struggle with basic math (see Appendix). They have problems with percentages, areas, and simple formulas (NAEP, 2010c). Their problems also appear in the international assessments (Gonzales et al., 2008, pp. B-5–B-8).

Test-prep helps explain why math performance has been poor and the gains fleeting—generally less at 8th grade than at 4th and mostly gone by high school. Drilling substitutes for teaching math in deeper ways that would lay a foundation for later work. This same pattern holds true in the international assessments. As they go through school, U.S. students fall farther behind those in other countries (cf. Gonzales et al., 2008, pp. 7, 9 and Baldi et al., 2007, pp. iv, 12).

In 2008, nearly twenty years after NCTM issued its standards, high school students were scoring the same in the long-term NAEP as they had been in the early 1990s (Rampey et al., 2009, p. 29). Had they really learned the math well at younger ages, or had a test-driven accountability system been the solution, their scores should have improved. Secondary school science performance, which depends on mathematical skills, also should have improved, yet it has stagnated nearly as long (Campbell et al., 2000, p. 9; Grigg, Lauko, & Brockway, 2006, pp. 19, 31). Such pervasive shortcomings tarnish the claims of success, even in mathematics.[2]

Judging Flat Trends

How should we judge generally flat national trends? From one perspective, maintaining scores is a sign of resilience given that schools have had to cope with changing social conditions. During the 1980s, single-parent households, teenage suicide, and juvenile arrests for violent crimes rose substantially (Stedman, 1998, p. 70). In the 1990s, the number of limited English proficiency students steadily increased (U.S. Department of Education, 2003, 2006).

Many changes over the past decades, however, were favorable and should have made the work of educators easier not harder. The population's educational attainment rose dramatically (NCES, 2010a, pp. 24–25). Pupil-teacher ratios declined substantially (pp. 60, 100)—as did student reported alcohol use (NCES, 2008, pp. 233–234). Even though child poverty and teen birth rates increased somewhat in recent years, both fell greatly in the 1990s and remain well below 1990 levels (Douglas-Hall & Koball, 2006, p. 3; Hamilton, Martin, & Ventura, 2009, p. 2; Martin et al., 2009, p. 2; NCES, 2008, pp. 35). Public school spending nearly quintupled during the past generation, rising from $118 billion in 1982–83 to over $600 billion by 2009 (NCES, 2010a, p. 48). Even adjusted for inflation, it has more than doubled.

Given these positive developments and the emphasis on standards and academics, achievement should have improved. Two other developments also should have pushed scores higher—a rising dropout rate during much of the standards era (discussed later) and the massive increase in attention to NAEP and its items as the program expanded to state and district testing and NCLB and states required schools to participate in its assessments (Stedman, 2009). State departments of education have set up web pages devoted to NAEP, provide guidance to schools participating in NAEP and, along with NAEP itself, have actively promoted the use of NAEP

[2] This is not the fault of NCTM's standards, which offered a needed alternative to traditional curriculum and pedagogy. Math reforms foundered, in part, on the shoals of a test-driven accountability system. On the other hand, the NCTM standards have been faulted for being based on flawed conceptions of mathematics and for a limited understanding of teacher work intensification and social and political inequality (Apple, 1992; Schmittau, 1991).

items in classroom instruction and assessment (California Department of Education, 2010; Foy, 2007, 2008; Georgia Department of Education, 2010; Illinois State Board of Education, 2010; Malone, 2010; NAEP, 2010a, 2010b; New Mexico Public Education Department, 2010).

Under such conditions, a tapestry of flat trends—now frayed by various international declines in math, science, and reading—is truly a poor result.

The Quality of Achievement

In the 1980s, the National Assessment Governing Board, which oversees NAEP, decided it would be useful to characterize student performance in terms of proficiency levels—basic, proficient, and advanced. Using panels of teachers, other educators, and members of the public, NAEP developed proficiency benchmarks, linking scale scores and items, and reported the percentages of students reaching each of those levels (Loomis & Bourque, 2001, pp. 2–3).

The impact on legislators, the media, and public debate was enormous. *Proficiency* became a watchword of state and federal policymakers. Goals 2000, passed in 1994, called for *all* students to be academically competent, defined as achieving proficiency in the main NAEP assessments (NEGP, 1995, pp. 11, 36). In turn, NCLB mandated that *all* students achieve proficiency in state assessments (NCLB, 2002).

Even after widespread and intense accountability efforts, however, the standards movement has fallen far short of this goal. In every area NAEP tests, most students, including most high school seniors, are still not proficient. On average, a third of high school seniors do not even reach the *basic* level, which reflects only a *partial* mastery of the fundamentals (Reese et al., 1997, p. 42) (see Table 1). The worst areas are science and U.S. history, where less than 20% of the students are proficient and about *half* do not even come up to the basic level.

Table 1

12th Grade Achievement in the 2001–2009 Main NAEP Assessments

Subject	Year	Below Basic	At Basic	Proficient or better	Advanced or better
Civics	2006	34%	39%	27%	5%
Economics	2006	21%	37%	42%	3%
Geography	2001	29%	46%	25%	1%
U.S. History	2006	53%	34%	13%	1%
Math	2009	36%	38%	26%	3%
Reading	2009	26%	36%	38%	5%
Science	2005	46%	36%	18%	2%
Writing	2007	18%	58%	24%	1%
Average		33%	41%	27%	3%

Note: "Proficient or better" includes students at the advanced level. Data are from the latest 12th grade assessments. Data sources: Lutkus & Weiss, 2007, p. 1 (civics); Mead & Sandene, 2007, p. 5 (economics); Weiss el al., 2002, p. 21 (geography); Lee & Weiss, 2007, p. 9 (U.S. history); NCES, 2010b, pp. 9, 26 (math and reading); Grigg, Lauko, & Brockway, 2006, p. 31 (science); and Salahu-Din et al., 2008, p. 37 (writing). The percentages in the average row do not total 100% due to rounding (e.g., 41% instead of 40.5%).

Achievement, however, is likely worse than these percentages indicate. The data do not include dropouts, which would lower the averages. The standards involved are weak ones: to be

judged proficient, students only have to answer correctly 65% of the constructed questions at the proficient level—or 74% of the multiple choice ones (Bourque & Garrison, 1991, pp. 3–6, Grigg et al., 2007, p. 13; Lutkus & Weiss, 2007, p. 31). A more reasonable standard, 80–90%, would mean even fewer students are proficient.

Although evaluators have questioned the validity of these levels, they still felt they could "be of use in describing changes in student performance over time" (Pellegrino, Jones, & Mitchell, 1999, p. 176). The standards movement, however, has hardly budged the percentages of high school students achieving proficiency—only 1-2 points in several subjects—and there have even been declines in others (see Appendix). The percentages of proficient students in reading, science, and geography are now slightly *less* than they were in the mid-1990s (NCES, 2010b, p. 9; Grigg, Lauko, & Brockway, 2006, p. 31; Weiss et al., 2002, p. 21).

Looking at test items bypasses the issues with the levels and shows that students struggle with basic material in each subject (see Appendix). High school students, for example, still lack basic knowledge of their nation's history and have weak civics and map skills. Most 12th graders do not know what the Lincoln-Douglas debates were about or why the U.S. was in the Korean War. Despite the Vietnam War and recurring Mideast crises, almost two-thirds in past assessments could not identify Southeast Asia, the Persian Gulf, or Saudi Arabia. There are major gaps in knowledge of race and constitutional issues, especially about school segregation and *Brown v. Board of Education*. Most high school students still have trouble with middle-school math and show striking deficiencies in earth science, chemistry, and biology. Their writing remains weak. The international assessments and other national surveys confirm these problems (Intercollegiate Studies Institute, 2007, 2008; National Science Board, 2008).

Learning also lacks depth. When researchers go past the standardized tests and probe understanding, they find students' knowledge of math, science, and social studies is superficial and that they misunderstand key concepts (Barker, 2004; Foster, 2007; Gonzalez-Espada, 2003; National Mathematics Advisory Panel, 2008; Schmittau, 2004; Wineburg, 1999). Adding to the worries, students show a disturbing lack of allegiance to core constitutional principles. For example, about half the country's high school students endorse government *censorship* of the news media (Knight Foundation, 2005).

In spite of a generation of efforts, therefore, academic performance remains poor. The conclusion of NAEP's 1990 long-term trend report still stands: "very few students demonstrate that they can use their minds well" (Mullis, Owen, & Phillips, 1990, p. 10).

Part II: The Transformation of the School

Academic Coursework

Besides achievement, the standards movement has emphasized academic coursework. In response to *A Nation at Risk,* states across the nation raised graduation requirements (Lillard & DeCicca, 2001, pp. 460–461; Planty, Provasnik, & Daniel, 2007, p. 2) and produced what looks like a staggering improvement. In 1982, only 10% of high school graduates had taken 4 years of English, 3 of social studies, math, and science, and 2 of a foreign language (NCES, 2008, p. 219). By 2005, *more than half* had done so. The national emphasis on math and science also apparently paid off, with credit gains in both areas (Planty et al., 2007, p. 20).

There are grounds for caution, however. The data do *not* include the 25% of students who

drop out, and have taken fewer academic courses. *General* math and science credits are counted. The data are also inflated by the inclusion of private school graduates, who take more academics. Even with that, high school graduates still fall short of traditional college-prep expectations. Half of them have taken two years or *less* of a foreign language—a percentage unchanged since 1990—and many have no credit in it at all (Planty et al., 2007, pp. 20, 26). About a fourth lack geometry and about a third are missing Algebra 2 (p. 21). Their overall science gain was just one course in over 20 years (p. 20). A third still have not taken chemistry, and about 40% graduate without having taken both biology and chemistry (NCES, 2008, p. 217; Planty et al., pp. 20–21). Only a third of the graduates have achieved *A Nation at Risk*'s New Basics graduation standards (NCEE, 1983; NCES, 2008, p. 219), while only 10% have taken what NCES labels a "rigorous" curriculum (Shettle et al., 2007, p. 10).

Many gains were concentrated in the 1980s and early 1990s. It was initially easy to enroll more students in geometry and chemistry, but progress then slowed. The 1980s gains also may not have reflected increased academic rigor. In a nationwide investigation, Toch (1991) found that schools had watered down existing courses, re-titled lower-level courses and created new ones to give the appearance of being academic, and permitted non-academic courses to substitute for graduation. He determined that

> The vast majority of the nation's high school students, as a result, are getting little more exposure to rigorous course work than they did previously. Despite the reformers' successful push for new graduation requirements, they are receiving an academic education in name only. (p. 102)

He concluded that the inference "increased graduation requirements are synonymous with higher academic standards is clearly false" (Toch, 1991, p. 105).

The value of academic courses decreased for another reason. During the past generation, the length of the school day and year changed little (Shettle et al., 2007, p. 7), yet schools crammed in more courses. Students once took about 5 courses a semester, a load many of us recall (Wright, Tomlinson, & Farris, 1985, p. 2). Today's graduates, however, average closer to 7 (NCES, 2008, pp. 211, 213). Such curricular cramming means each subject now receives less attention and—with shorter class periods—there is also less time for quality, in-depth lessons.

The AP Program

The standards movement also fueled the growth of the Advanced Placement program, yet many observers worry it has expanded too rapidly and that schools—and parents—are pushing unprepared students into courses of dubious quality (Cech 2007, 2008; Humes, 2003, p. 349; Matthews, 2008). Even AP teachers report declining student quality and that AP enrollments are being used by students to pad their college records—and by administrators to promote their school's image (Duffett & Farkas, 2009). Few of the teachers believe that the "earlier grades are doing a better job of preparing students for the rigor of AP coursework" (p. 1).

Indeed, AP scores have been declining and the percentage scoring 3 or higher, usually required for college credit, has fallen (College Board, 2009). Given the perceived loss of quality of the AP, some colleges and universities are now requiring scores of 4 or 5 or are refusing to grant credit altogether (Humes, 2003, p. 349). The problem became so serious that the College Board launched a nationwide audit to determine if schools were watering down the courses

(Cech, 2007). However, the audit was roundly criticized; it was of course syllabi only—not actual course content, instruction, or student performance—and teachers and schools were given extensive help in revising their syllabi and multiple chances to resubmit them for approval. Nevertheless, in the aftermath of the audit, over 2,000 schools stopped offering AP courses—either voluntarily or because their courses had been rejected by the College Board.

Even when courses are strong, however, exam prep takes its toll on students. As Humes (2003, p. 114) reports,

> for many, the joy—and more importantly, the ability to retain and continue to use what they memorize for the test—can be sucked from the AP experience: Kids are learning to pass a test on French or biology or civics, but their interest in the subject may go no further, or may even be extinguished, by the rigors of AP.

In any case, the AP program remains the province of an elite and not a bellwether of the standards movement. Most students do not take AP courses and a third of public high schools do not even offer them (NCES, 2008, pp. 66, 220; Planty et al., 2007, p. 18).

We should be skeptical, therefore, about increased academic enrollments as a measure of success, especially as they have not improved high school achievement. One could even view them more as a form of social promotion than as an indicator of academic rigor and improvement.[3]

Changes in School Culture

Educational Taylorism

Beyond achievement and coursework, we need to consider the effects the standards movement is having on the lives of students and teachers. There is growing evidence that, in its press to raise test scores, the movement has damaged teaching and stultified classrooms. These problems show up in case studies, state and national surveys, and NCLB assessments. In what follows, I discuss the findings from these diverse sources, including studies of urban schools and two highly-regarded suburban high schools that vividly illustrate what is going on.

I begin with McNeil's (2000) study of the Texas accountability system's impact on Houston's schools. There was no "Texas miracle." Using classroom observations, interviews with teachers and students, and investigations of historical developments, McNeil describes how the state-imposed system of "proficiency curricula" and "reform controls" devastated "authentic, engaged teaching and learning," treated teachers like employees rather than professionals, and worsened inequities (pp. 3–6, 154, 201, 204). Even in highly successful magnet schools, classroom teaching had "shifted away from intellectual activity toward dispensing packaged

---

[3] This is *not* an argument against students taking intellectually challenging courses, especially when combined with caring, learner-friendly environments. Recently, Toch and colleagues offered a positive, yet cautionary, report on school reform projects that combine rigorous college-prep courses and personalized instruction within small communities (Toch, Jerald, & Dillon, 2008). As we shall see in the next section, however, such environments are hardly the goal or result of test-driven, accountability interventions. Nor is their existence or worth reflected in the sheer number of academic course credits being accumulated. Quantity does not, *per se*, indicate quality.

fragments of information sent from an upper level of the bureaucracy" (p. 5).

Overall, McNeil found that the system had bred "a cycle of lowered expectations" (p. 154) and that the "panacea" offered by the testing-accountability system had "*two great costs: it undermines quality, and it increases discrimination*" (p. xxii, emphasis original). She concluded that the failures of "legislated learning" in Texas offer

> a powerful lesson for a nation increasingly preoccupied with calls for "standards" and for narrow measurement indicators as keys to holding its public schools "accountable" to those who would control them. They provide a powerful warning about effects on teaching and learning when all the authority over all the significant decisions about teaching and learning is centralized and all the means are standardized. (p. 154)

We see these effects in cities across the country, especially in schools serving minority students. In *The Shame of the Nation,* Kozol (2005) investigated schools in 11 states and 30 districts, including New York, Chicago, Los Angeles, and Milwaukee. He found pervasive segregation and offered a chilling portrait of urban education devoted to scripted instruction, "Skinnerian approaches," and the "pedagogy of direct command" (p. 64). Administrators have become obsessed with rules and standardization. There is a "fanatical insistence" on uniformity and time management (p. 64).[4]

As part of the mania for control, schools have instituted silent lunches and even imposed silent recesses. Many have dropped recess all together (Blackwell, 2004; Ohanian, 2002; Patte, n.d.). The time-on-task rationales are revealing. Atlanta's superintendent argued, "We are intent on improving academic performance. You don't do that by having kids hanging on the monkey bars" (Johnson, 1998; Kozol, 2005, p. 120). Yet, fresh air and regular exercise can help students do well academically and are especially needed in an era of children's growing obesity and fitness problems.[5]

Curricular Constriction

Schools should nurture students aesthetically as well. Yet, across the country, principals and teachers report that time on the arts is being reduced and that instruction is increasingly test-focused (Orfield, 2006, p. 7; von Zastrow & Janc, 2004, pp. 7–8). NCLB is a prime contributor

---

[4] It is fashionable in some academic circles to dismiss Kozol's work as "journalistic," but over the years, when our research has coincided, e.g., on illiteracy and urban education, I have found him a skilled researcher and astute observer. He has marshaled well research findings, educational history and data, and legal decisions. His portraits of urban schools and children have the quality of fine, critical-theory influenced, case study research. Even in the terms of the disparagement, he is no mere journalist, but a skilled investigative reporter and researcher. That is of great value when juxtaposed against corporate-dominated media that serve elites and help perpetuate social injustice.

[5] The rote pedagogy and restrictive environments are reminiscent of the harsh practices of 19th century schooling. An 1890s mindset chains students to their desks, fills their days with test preparation, and neglects the whole child. In the early 20th century, progressive educators and the playground movement challenged such severe approaches and brought outdoor spaces even to congested city schools (Cavallo, 1981; Frost & Wortham, 1988; Reese, 2002). Educators came to recognize the benefits of outdoor activity and exercise for proper development and improved academics. The abandonment of recess represents a disturbing and unwise retreat to an earlier era.

to these trends (Wood, 2004, pp. 34–36). In a series of nationally representative studies, the Center on Education Policy found that school districts have refocused their curricula on the material covered by NCLB-mandated tests (McMurrer, 2007, pp. 1–2) (see Appendix). They have substantially increased time on reading and math, while reducing it in social studies and the arts (McMurrer, 2007, p. 8; McMurrer, 2008, pp. 2–6).

The conception of good education is narrowing as well. In a mixed-methods analysis of California middle schools, Mintrop and Trujillo (2007) found that schools that had improved test scores in response to the state's aggressive benchmarking system had embraced "a rather constricted notion of quality," "one that excludes, for the most part, quality of teaching" (p. 37). What teachers did to raise scores was not synonymous with "higher academic engagement of their students, better teaching, or more learning complexity" (p. 35). Mintrop and Trujillo wryly noted that "better aligned" schools are not the same thing as "better" schools (p. 37).

This should not be surprising. When institutions are required to improve outcomes, they invariably focus on what is being measured and rewarded, even if it means short-changing the things most people value. We have been through this before. Schools that adopted the effective schools formula of frequent testing and basic skills often sacrificed critical thinking skills and the development of aesthetic sensibilities and an ethic of sharing (Cuban, 1983; Stedman, 1985). When the international assessments showed Japan doing well and the U.S. doing poorly, the neoconservative clamor was to copy Japan's national testing program and its longer school year, but that ignored its constructivist pedagogy, frequent recess breaks, and emphasis on the arts.

The Price of "Success"

The damage to success-driven suburban students has taken a different form. In *School of Dreams*, Pulitzer-prize winner Humes (2003) documents the price being paid at a top California high school. Anxiety-ridden students obsessively pursue extracurricular activities to pad their records. Marijuana, Starbucks, and alcohol are a fixture in many lives. Cheating and plagiarism are serious problems. Students have become "test-taking machines" who work "until they score the grade they need, then move on" (pp. 94–95). They are unresponsive even to imaginative teaching, and "too ready to accept official pronouncements" (p. 72). This is especially true in AP classes, which are devoted to exam prep and "more rigidly programmed" (p. 72).

In its past, the high school had been an intellectual oasis where "young, excited, interesting" teachers taught "interdisciplinary core classes" (Humes, 2003, p. 25). It had an "intimacy and unexpected sense of family" (p. 26). Students were internally motivated and worked hard. This special place succumbed, however, to the pressures of the standards era and became a "far more stressful, far more caffeinated incarnation" (p. 188). Humes reported that a recently-deceased veteran teacher was much beloved because she was one of those who still

> believed in the importance of a kind of learning that could not be measured by SATs or GPAs or the ever-growing legion of standardized tests mandated by state and federal governments, tests that have everything to do with satisfying bureaucracies but which, at best, quantify only a piece of what ought to go on in a classroom. (p. 93)

Rituals of Compliance: "Doing School"

What Humes found has been witnessed by others, especially in schools where testing and

"excellence" have taken center stage. In a rich study, with the counsel of Eisner, Peshkin, and McDermott, Clark-Pope (2001) shadowed several of the best and brightest at a successful college-preparatory high school. The school has a "reputation for caring teachers, innovative programs, and strong leadership" (Clark-Pope, p. xiv). The students "seem to be diligent, talented, and focused"; they "get good grades, win awards and commendations, pursue extracurricular interests, do community service, and help teachers and administrators" (p. 149). In short, they are just the type of students standards advocates claim they would like to produce. Yet, their lived reality is in stark contrast to the glowing outcomes. To succeed, the students lie, cheat, plagiarize, skip class, compromise their values, and suffer from health problems, including sleeplessness and anxiety.

Clark-Pope links the problems to features of the modern high school: competitiveness, regimented time structures, overworked teachers, fragmented curricula, and the grade trap. She finds that the grade-driven system provides "little support for intellectual engagement and passion" (p. 155) and instead rewards memorization, deception, and shallow understanding. Clark-Pope concludes successful students have mastered "doing school"; that is, being adept at taking shortcuts, manipulating adults, and outwitting the bureaucracy. The irony is that by "overemphasizing achievement," the standards movement has sabotaged learning (Kohn, 2004).

We should not glorify high schools of the past, however. From the 1950s to the 1980s, authoritarian high schools alienated students, undermined academic engagement, and distributed knowledge unequally (Goodlad, 1983; Silberman, 1970; Sizer, 1984). Yet, many parents and teachers accepted the arrangement in exchange for students' easy acquisition of credentials. Since then, testing mandates have taken over the schools' educative function and the focus is on meeting external accountability demands, no matter how unfulfilling or intellectually bereft they are. The treaties of mediocrity that once plagued schools have been replaced by rituals of compliance. Dewey's (1938, pp. 26–27) scathing critique of traditional education still applies:

> the experiences which were had, by pupils and teachers alike, were largely of a wrong kind. How many students, for example, were rendered callous to ideas, and how many lost the impetus to learn because of the way in which learning was experienced by them? How many acquired special skills by means of automatic drill so that their power of judgment and capacity to act intelligently in new situations was limited?

Heading for the Exits

The standards movement may aim for students to achieve well and graduate from high school but, by creating a school culture antithetical to student engagement and genuine teaching and learning, it has helped raise dropout rates, increase teacher burnout, and short-circuit reading.

Student Dropouts

Recently, a Nobel-prize winning economist and his colleague determined that graduation rates are inflated and that the true rate "peaked in the early 70s" and then "steadily declined" (Heckman & LaFontaine, 2007, p. 17). In 1984, the year after *A Nation at Risk* was released, nearly four-fifths of public school students were graduating from high school, but by 2001, only three-fourths were (see Appendix). Minority students lag far behind, and the gap has "not

converged over the past 35 years" (Heckman & LaFontaine, 2007, p. 3). Other economists found that, net of other factors, as states increased graduation requirements, dropouts increased from 3-8% (Lillard & DeCicca, 2001; Olson, 2000). This is one more reason high school test scores should have improved rather than stagnated. It means that flat NAEP trends and the U.S. declines in the international assessments are even worse than they appear.

Calling the dropout problem the "silent epidemic," the Gates Foundation reported an even lower graduation rate, closer to two-thirds, with a minority graduation rate of only half (Bridgeland, DiIulio, & Morison, 2006, p. i). It found that the situation has "not substantially improved during the past few decades when education reform has been high on the public agenda" (p. i). It observed further that, "During this time, the public has been almost entirely unaware of the severity of the dropout problem due to inaccurate data" (Bridgeland, DiIulio, & Morison, 2006, p. i). Pressures on school officials have led to inflated graduation rates and even accounting scandals.

After NCLB was enacted, the graduation rate briefly jumped, but economists surmised its mandated reporting gives "schools strong incentives to raise graduation rates by any means possible" (Heckman & LaFontaine, 2007, p. 17). Many states now maintain two sets of books (Dillon, 2008). Mississippi reported an 87% graduation rate for NCLB, when its actual rate was only 63%. California reported 83% to the federal government, but its rate was 67%. New Mexico counted only 12th graders who had not graduated, thus slyly ignoring most who leave school earlier. The discrepancies became so bad that Secretary of Education Spellings promulgated new accounting rules requiring districts to track cohorts. In any event, the graduation rate "never fully recovered to its early levels" (Heckman & LaFontaine, 2007, p. 17) and has now leveled off (NCES, 2010a, p. 164).

The Gates Foundation also investigated why students drop out. For all of the standards effort, nearly half felt poorly prepared for high school. Half reported their schools were boring, while nearly 70% said they "were not motivated or inspired to work hard" (Bridgeland et al., 2006, p. iii). Johns Hopkins researchers recently identified many high schools as "dropout factories" (Balfanz, n.d.; Gibson, 2008; Zuckerbrod, 2007). Is it any wonder critical theorists and many educators describe the problem as one of "push-outs" rather than "dropouts"?

Teacher Dropouts

Another devastating development is the skyrocketing *teacher* dropout rate. It is up 50% in the last two decades, and over 8% of public school teachers now leave annually (Marvel et al., 2007, p. 7). The national Teacher Follow-up Survey found that many who recently left public schools were dissatisfied with teaching as a career, about a third retired, and about a fourth sought a job other than K–12 teaching (p. 14). Those who ended up outside of education reported better working conditions, greater recognition and support, and a more manageable workload (p. 16). By a vast margin, they felt they now had better "influence over workplace policies and practices." Teaching was favored in only two of 20 areas: employment benefits and the opportunities to make a difference in the lives of others.

While the standards movement is not solely responsible for driving out teachers, it is contributing. Intrusive accountability systems have intensified teachers' work, while externally imposed curricula have deskilled teachers and subverted professionalism (Apple, 1992; McNeil, 2000; Ross, 1996). NCLB is directly implicated. In a recent national survey, most teachers report

it is promoting teaching-to-the-test, narrowing the curriculum, and increasing teacher burnout (Teachers Network, 2007a, 2007b). Members of the International Reading Association and the Association of Supervision and Curriculum Development have reached similar judgments (Fair Test, 2007; Roller, n.d.). Teachers deciding to leave, however, should not be our only concern. Overworked, burned-out teachers, who remain in schools, are more likely to use routine instruction and alienate students (Haberman, 2004).

Growing Aliteracy

Test-driven instruction and narrow, basic-skills curricula may also have taken a toll on reading. For many students, reading has become a chore associated with slogging through homework and cramming for tests. It certainly has lost much of its appeal as a leisure activity. In 2004, a third of U.S. 17-year-olds reported they *never* read on their own or did so only a few times a year (Perie et al., 2005, p. 55). This rate had nearly doubled in 20 years. As students go through school, they become less interested in reading. While a majority of 9-year-olds report reading for fun regularly ("almost every day"), only about a fifth of high school students do. Decades ago, Dewey (1938, p. 27) captured well the detrimental impact of drill-based schooling on students, "How many came to associate books with dull drudgery, so that they were 'conditioned' to all but flashy reading matter?"

The long-term effects are already showing up. In *Reading at Risk*, the National Endowment for the Arts documented major declines in adults' leisure reading of "novels, short stories, plays, or poetry" between 1982 and 2002 (NEA, 2004, p. ix). More than half *never* read any literature. While NEA's chairman blamed the problem on the growing use of electronic media, test-prep pedagogy has likely reinforced the trends. The shallow text processing involved in skimming dumbed-down textbooks and riffling through pre-packaged curriculum materials mirrors that of cell-phone texting and web surfing.

In its research review, the NEA (2007) reported that both literary *and* non-literary reading had declined. Making this worse, the declines were especially noticeable among college graduates and young adults. While fiction reading has increased a bit recently, the percentage of adults who *never* read books outside of school and work has continued to grow (NEA, 2009, p. 7; Thompson, 2009). Without passionately engaging students in literary activities, however, especially book reading, there is little hope of countering these trends. Infrequent and superficial reading helps explain why students' learning lacks depth and why they often do so poorly in the national and international assessments. Sadly, though predictably, aliteracy has become a major problem.

The Failure of the Standards Movement

After more than 25 years of vigorous effort, from *A Nation at Risk* and Goals 2000 to state testing and NCLB, the lack of progress is stunning. In spite of systematic changes in curricular standards and graduation requirements and the imposition of state and national accountability systems, profound achievement problems persist and the minority achievement gaps remain vast. High school performance continues to stagnate. While the standards movement did not create these deep-seated problems, it has been unable to fix them. At the same time, standards-based accountability systems have contributed to making things worse, not better, by warping school culture and adversely affecting the lives of students and teachers. Dropout rates increased for students and went up dramatically for teachers. We need a radically new approach to schooling, but to construct it well, it is important to first understand the reasons for the

movement's failure.

The second paper, "Why the Standards Movement Failed," is devoted to those matters. In it, I discuss the educational logic and political forces behind the standards movement. I examine the movement's premises and its connections to the neoliberal reform project and capitalism. I describe the root problems of NCLB and Race to the Top and propose a progressive alternative for school reform.

References

Apple, M. (1992). Do the standards go far enough? *Journal for Research in Mathematics Education*, *23*(5), 412–431.

Applebee, A., Langer, J., & Mullis, I. (1987). *Literature & U.S. history: The instructional experience and factual knowledge of high school juniors*. Princeton, New Jersey: ETS.

Baker, D. (1993). Compared to Japan, the U.S. is a low achiever…Really. *Educational Researcher, 22*(3), 18–20.

Baldi, S., Jin, Y., Skemer, M., Green, P., Herget, D., & Xie, H. (2007). *Highlights from PISA 2006: Performance of U.S. 15-year-old students in science and mathematics literacy in an international context*. Washington, DC: U.S. Department of Education.

Balfanz, R. (n.d.). *More information on the methodology, data, and terms used in the AP dropout factory story.* Center for Social Organization of Schools at Johns Hopkins University. Baltimore, MD: Johns Hopkins University. Retrieved from http://web.jhu.edu/CSOS/images/AP.html

Barker, V. (n.d.). *Beyond appearances: Students' misconceptions about basic chemical ideas*. A report prepared for the Royal Society of Chemistry. Retrieved from http://www2.iq.usp.br/docente/famaxim/disciplina/qfl3502/Students_misconceptions_about_basic_chemical_ideas.pdf (Note: At one point there was a 2004 2nd edition by V. Kind.)

Blackwell, J. (2004, August 6). Recess: Forgotten, neglected, crossed off, or hidden. *Childhood Education*. Retrieved from http://findarticles.com/p/articles/mi_qa3614/is_200401/ai_n9356956

Bourque, M., & Garrison, H. (1991). *The LEVELS of mathematics achievement*. Washington, DC: National Assessment Governing Board.

Braswell, J., Lutkus, A., Grigg, W., Santapau, S., Tay-Lim, B., & Johnson, M. (2001). *The nation's report card: Mathematics 2000*. Washington, DC: G.P.O.

Bridgeland, J., DiIulio, Jr., J., & Morison K. (2006). *The silent epidemic: Perspectives of high school dropouts*. Washington, DC: Civic Enterprises.

Bronner, E. (1998, February 25). U.S. trails the world in math and science. *New York Times*.

Bush, G. (2008). *No Child Left Behind in the new year*. Retrieved from http://www.presidentialrhetoric.com/speeches/01.07.08.html

California Department of Education. (2010). *National Assessment of Educational Progress (NAEP)*. Retrieved from http://www.cde.ca.gov/ta/tg/nr/

Camara, W. & Powers, D. (1999, July). Coaching and the SAT I. *The Industrial-Organizational Psychologist*, *37*(1). Retrieved from http://www.siop.org/tip/backissues/TipJuly99/4Camara.aspx

Campbell, J., Hombo, C., & Mazzeo, J. (2000). *NAEP 1999 Trends in academic progress: Three decades of student performance*. Washington, DC: Government Printing Office.

Campbell, J., Reese, C., O'Sullivan, C., & Dossey, J. (1996). *NAEP 1994 trends in academic progress*. Washington, DC: U.S. Department of Education.

Campbell, J., Voelkl, K., & Donahue, P. (1998). *NAEP 1996 trends in academic progress. Addendum. Achievement of U.S. students in science, 1969 to 1996; mathematics, 1973 to 1996; reading, 1971 to 1996; writing, 1984 to 1996. Revised.* Publication No. 98-531. Washington, DC: NCES.

Carpenter, T., Lindquist, M., Brown, C., Kouba, V., Silver, E., & Swafford, J. (1988). Results of the fourth NAEP assessment of mathematics. *Arithmetic Teacher* (December), 38–41.

Cataldi, E., Laird, J., KewalRamani, A., & Chapman, C. (2009). *High school dropout and completion rates in the United States: 2007 compendium report.* Washington, DC: NCES.

Cavallo, D. (1981). *Muscles and morals: Organized playgrounds and urban reform, 1880–1920.* Philadelphia: University of Pennsylvania Press.

Cech, S. (2007, November 14). Number of schools offering AP falls after first audit of course: Educators divided over first-ever audit of AP courses. *Education Week*, *27*(12), 1, 13.

Cech, S. (2008, February 20). AP trends: Tests soar, scores slip. *Education Week*, *27*(24), 1, 13.

Center on Education Policy. (2006). *From the capital to the classroom: Year 4 of the No Child Left Behind Act*. Washington, DC: Author.

Center on Education Policy. (2007, July 25). As the majority of school districts spend more time on reading and math, many cut time in other areas. Press release. Retrieved from http://www.cep-dc.org/press/Curriculum%20Release%20Final.pdf

Cheney, L. (n.d.). *American memory: A report on the humanities in the nation's public schools*. Washington, DC: U.S. Government Printing Office.

Clark-Pope, D. (2001). *Doing school: How we are creating a generation of stressed out, materialistic, and miseducated students*. New Haven: Yale University Press.

College Board, The. (2009). *AP grade distributions: All subjects 1988–2008.* New York: The College Board. Retrieved from http://professionals.collegeboard.com/profdownload/ap-data-2008-Grade-Dist-All.pdf

College Board, The. (2010). *College-bound seniors 2010.* New York: The College Board. http://professionals.collegeboard.com/profdownload/2010-total-group-profile-report-cbs.pdf

Cuban, L. (1983). Effective schools: A friendly but cautionary note. *Phi Delta Kappan*, *64*(10), 695–96.

Dewey, J. (1938). *Experience and education.* New York: Macmillan Company.

Dillon, S. (2008, March 20). States' data obscure how few finish high school. *New York Times*. Retrieved from http://www.nytimes.com/2008/03/20/education/20graduation.html

Dillon, S. (2009, February 22). Rename law? No wisecrack is left behind, *New York Times*. Retrieved from http://www.nytimes.com/2009/02/23/education/23child.html

Douglas-Hall, A., & Koball, H. (2006). *The new poor: Regional trends in child poverty since 2000.* Retrieved from http://www.nccp.org/publications/pdf/text_672.pdf

Duffett, A., & Farkas, S. (2009). *Growing pains in the Advanced Placement program: Do tough trade-offs lie ahead?* Washington, DC: Fordham Institute. Retrieved from http://edexcellence.net/index.cfm/news_advanced-placement-program-study

Elley, W. (1992). *How in the world do students read?* Hamburg: International Association for the Evaluation of Educational Achievement.

Fair Test (2007, April). Polls: Teachers reject NCLB. *Fair Test Examiner*. Retrieved from http://www.fairtest.org/polls-teachers-reject-nclb

Foster, D. (2007). Making meaning in algebra: Examining students' understandings and misconceptions. In A. Schoenfeld (Ed.), *Assessing mathematical proficiency* (Chapter 12, pp. 163–177). New York: Cambridge University Press.

Foy, J. (2007). *NAEP assessment toolbox for teachers. Easy-to-use classroom activities using questions and data from the National Assessment of Educational Progress*. Alaska NAEP Coordinator. Alaska Department of Education and Early Development. Retrieved from http://www.eed.state.ak.us/tls/assessment/naep/Fall07/NAEPQAssessmentToolboxforTeachers.doc

Foy, J. (2008). *Using NAEP assessment items instructionally: Teaching students to be strategic*. Alaska NAEP Coordinator. Alaska Department of Education and Early Development. Retrieved from http://www.eed.state.ak.us/tls/assessment/naep/Fall08/UsingNAEPItems.ppt

Frost, J., & Wortham, S. (1988). The evolution of American playgrounds. *Young Children*, *43*(5), 19–28.

Fuller, B., Wright, J., Gesicki, K., & Kang, E. (2007). Gauging growth: How to judge No Child Left Behind?, *Educational Researcher*, *36*(5), 268–278.

Georgia Department of Education. (2010). *National Assessment of Educational Progress*. Retrieved from http://doe.k12.ga.us/ci_testing.aspx?PageReq=CI_TESTING_NAEP

Gibson, R. (2008, August). *Chicago, Detroit, schools, and the election spectacle*. Retrieved from http://www.richgibson.com/chicago-detroit-schools-election.htm

Glod, M. (2009, April 29). "Nation's report card" sees gains in elementary, middle schools. *Washington Post*. Retrieved from http://www.washingtonpost.com/wp-dyn/content/article/2009/04/28/AR2009042801244.html?hpid=moreheadlines

Gonzales, P., Williams, T., Jocelyn, L., Roey, S., Kastberg, D., & Brenwald, S. (2008). *Highlights from TIMSS 2007: Mathematics and science achievement of U.S. fourth- and eighth-grade students in an international context*. NCES 2009-001. Washington, DC: U.S. Department of Education.

Gonzalez-Espada, W. (2003). A last chance for getting it right: Addressing alternative conceptions in the physical sciences. *The Physics Teacher*, *41*(1), 36–38.

Goodlad, J. (1983). *A place called school.* New York: McGraw-Hill.

Greenwald, E., Persky, H., Campbell, J., & Mazzeo, J., with Jenkins, J. & Kaplan, B. (1999). *The NAEP 1998 writing report card for the nation and the states.* Washington, DC: U.S. Department of Education.

Grigg, W., Donahue, P., & Dion, G. (2007). *The nation's report card: 12th grade reading and mathematics 2005*. Washington, DC: U.S. Government Printing Office.

Grigg, W., Lauko, M., & Brockway, D. (2006). *The nation's report card: Science 2005*. NCES 2006-466. Washington, DC: U.S. Government Printing Office.

Haberman, M. (2004). Teacher burnout in black and white. *Education News*, January 9, 2004. Retrieved from http://www.ednews.org/articles/teacher-burnout-in-black-and-white.html

Hamilton, B., Martin, J., & Ventura, S. (2009). Births: Preliminary data for 2007. *National Vital Statistics Reports*, *57*(12). Web release. Hyattsville, MD: National Center for Health Statistics. http://www.cdc.gov/nchs/data/nvsr/nvsr57/nvsr57_12.pdf

Heckman J., & LaFontaine, P. (2007). *The American high school graduation rate: Trends and levels*. Discussion Paper No. 3216. Bonn, German: Institute for the Study of Labor.

Hoff, D. (2007, June 20). State tests show gains since NCLB: Report cautions against crediting education law. *Education Week*, *26*(39), 1, 20.

Humes, E. (2003). *School of dreams: Making the grade at a top American high school*. New York: Harcourt.

Illinois State Board of Education. (2010). *State assessment: National Assessment of Educational Progress (NAEP*). Retrieved from http://www.isbe.state.il.us/assessment/naep.htm

Improving America's Schools Act of 1994, Pub. L. No. 103-382. (1994). Retrieved from http://www2.ed.gov/legislation/ESEA/toc.html Accountability provisions retrieved from http://www.ed.gov/legislation/ESEA/sec1111.html

Intercollegiate Studies Institute. (2007). *Failing our students, failing America: Holding colleges accountable for teaching America's history and institutions*. A report by the Intercollegiate Studies Institute's National Civic Literacy Board. Wilmington, DE: ISI.

Intercollegiate Studies Institute. (2008). *Our fading heritage: Americans fails a basic test on their history and institutions.* Wilmington, DE: ISI.

Johnson, D. (1998, April 7). Many schools putting an end to child's play. *New York Times.*

Kaufman, P., Alt, M., & Chapman, C. (2004). *Dropout rates in the United States: 2001*. Washington, DC: U.S. Government Printing Office.

Kendall, J. & Marzano, R. (2004). *Content Knowledge*: *A compendium of standards and benchmarks for K-12 education*. 4th edition. Aurora, CO: Mid-continent Research for Education and Learning (McREL). (Information on the funding and history of standards documents is available at http://www.mcrel.org/standards-benchmarks/docs/purpose.asp and http://www.mcrel.org/standards-benchmarks/docs/history.asp)

Knight Foundation (2005). *The future of the first amendment: What America's high school students think about their freedoms*. Miami, Florida: Knight Foundation.

Kohn, A. (2004). The costs of over-emphasizing achievement. In S. Mathison & E. W. Ross (Eds.), *Defending public schools, Volume IV: The nature and limits of standards based reform and assessment* (Chapter 3, pp. 27–34). Westport, Connecticut: Praeger.

Kozol, J. (2005). *The shame of the nation: The restoration of apartheid schooling in America*. New York: Three Rivers Press.

Lee, J. (2006). *Tracking achievement gaps and assessing the impact of NCLB on the gaps*. Cambridge, MA: The Civil Rights Project at Harvard University.

Lee, J. & Weiss, A. (2007). *The nation's report card: U.S. history 2006.* NCES 2007–474. Washington, DC: U.S. Government Printing Office.

Lemke, M., Sen, A., Pahlke, E., Partelow, L., Miller, D., Williams, T., Kastberg, D., & Jocelyn, L. (2004). *International outcomes of learning in mathematics literacy and problem solving: PISA 2003 results from the U.S. perspective*. NCES 2005–003. Washington, DC: U.S. Department of Education.

Lillard, D. & DeCicca, P. (2001). Higher standards, more dropouts? Evidence within and across time. *Economics of Education Review*, *20*(5), 459–473.

Loomis, S. & Bourque, M. (Eds.) (2001). *National Assessment of Educational Progress achievement levels 1992–1998 for civics*. Washington, DC: U.S. Department of Education.

Lutkus, A. & Weiss, A. (2007). *The nation's report card: Civics 2006.* NCES 2007–476. Washington, DC: U.S. Government Printing Office.

Malone, J. (2010). *The nation's report card*. PowerPoint document by the Mississippi NAEP State Coordinator. Mississippi Department of Education. Retrieved from http://www.mde.k12.ms.us/ACAD/osa/DTC%20Feb%202010/PP%20NAEP%20Feb%20DTC%202010.ppt

Martin, J., Hamilton, B., Sutton, P., Ventura, S., Menacker, F., Kirmeyer, S., & Mathews, T. (2009). Births: Final data for 2006. *National Vital Statistics Reports*, *57*(7). Hyattsville, MD: National Center for Health Statistics. Retrieved from http://www.cdc.gov/nchs/data/nvsr/nvsr57/nvsr57_07.pdf

Martin, M., Mullis, I., Gonzalez, E., & Chrostowski, S. (2004). *TIMSS 2003 international science report*. Chestnut Hill, MA: Boston College.

Martin, M., Mullis, I., Gonzalez, E., & Kennedy, A. (2003). *Trends in children's reading literacy achievement 1991–2001: IEA's repeat in nine countries of the 1991 reading literacy study*. Chestnut Hill, MA: Boston College.

Marvel, J., Lyter, D., Peltola, P., Strizek, G., Morton, B., & Rowland, R. (2007). *Teacher attrition and mobility: Results from the 2004–05 teacher follow-up survey: First look*. Washington, DC: U.S. Government Printing Office.

Matthews J. (2008, March 10). Should we put the brakes on advanced placement growth? *Washington Post*. Retrieved from http://www.washingtonpost.com/wp-dyn/content/article/2008/03/10/AR2008031000401.html

McMurrer, J. (2007). *Choices, changes, and challenges: Curriculum and instruction in the NCLB era*. Washington, DC: Center on Education Policy.

McMurrer, J. (2008). *Instructional time in elementary schools: A closer look at changes for specific subjects*. Washington, DC: Center on Education Policy.

McNeil, L. (2000). *Contradictions of school reform: Educational costs of standardized testing*. New York: Routledge.

Mead, N., & Sandene, B. (2007). *The nation's report card: Economics 2006*. NCES 2007–475. Washington, DC: U.S. Government Printing Office.

Mintrop, H. & Trujillo, T. (2007). *The practical relevance of accountability systems for school improvement: A descriptive analysis of California schools.* Los Angeles: National Center for Research on Evaluation, Standards, and Student Testing.

Mullis, I., Martin, M., Gonzalez, E., & Chrostowski, S. (2004). *TIMSS 2003 international mathematics report*. Chestnut Hill, MA: Boston College.

Mullis, I., Martin, M., Kennedy, A., & Foy, P. (2007). *IEA's progress in international reading literacy study in primary school in 40 countries*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

Mullis, I., Owen, E., & Phillips, G. (1990). *America's challenge: Accelerating academic achievement. A summary of findings from 20 years of NAEP*. Report No. OV-01. Princeton, New Jersey: Educational Testing Service.

National Assessment Governing Board. (2008). *Science framework for the 2009 National Assessment of Educational Progress.* Washington, DC: U.S. Department of Education.

National Assessment of Educational Progress. (2010a). *2011 facts for principals*. Retrieved from http://nces.ed.gov/nationsreportcard/pdf/about/schools/2011PrincipalsFactSheet.pdf

National Assessment of Educational Progress. (2010b). *An introduction to NAEP: National Assessment of Educational Progress*. Retrieved from http://nces.ed.gov/nationsreportcard/pdf/parents/2010468.pdf

National Assessment of Educational Progress. (2010c). *NAEP questions.* Test items from on-line database Retrieved from http://nces.ed.gov/nationsreportcard/itmrls/startsearch.asp

National Center for Education Statistics. (1993). *Digest of education statistics 1993*. U.S. Department of Education. Washington, DC: U.S. Government Printing Office.

National Center for Education Statistics. (2008). *Digest of education statistics 2007*. U.S. Department of Education. Washington, DC: U.S. Government Printing Office.

National Center for Education Statistics. (2009). *The nation's report card: Mathematics 2009* (NCES 2010–451). Washington, DC: U.S. Government Printing Office.

National Center for Education Statistics. (2010a). *Digest of education statistics 2009*. U.S. Department of Education. Washington, DC: U.S. Government Printing Office.

National Center for Education Statistics. (2010b). *The nation's report card: Grade 12 reading and mathematics 2009 national and pilot state results*. (NCES 2011–455). Washington, DC: U.S. Department of Education.

National Center for Education Statistics. (2010c). *The nation's report card: Reading 2009* (NCES 2010–458). Washington, DC: U.S. Government Printing Office.

National Commission on Excellence in Education. (1983). *A nation at risk.* Washington, DC: Government Printing Office.

National Council of Teachers of Mathematics. (1989). *Curriculum and evaluation standards for school mathematics*. Reston, VA: Author.

National Council of Teachers of Mathematics. (2004, January/February). Rise in NAEP math scores coincides with NCTM standards. *News Bulletin*. Retrieved from http://www.nctm.org/news/release.aspx?id=766

National Council of Teachers of Mathematics. (2005, July 14). Trend NAEP reports math scores up for 9- and 13-year-olds; 17-year-olds steady. Press release. Retrieved from http://www.nctm.org/news/content.aspx?id=714

National Education Goals Panel. (1995). *The national education goals report: Building a nation of learners 1995*. Washington, DC: U.S. Government Printing Office.

National Education Goals Panel. (1999). *The national education goals report: Building a nation of learners 1999*. Washington, DC: U.S. Government Printing Office.

National Endowment for the Arts. (2004). *Reading at risk: A survey of literary reading in America*. Research Report #46. Washington, DC: NEA.

National Endowment for the Arts. (2007). *To read or not to read: A question of national consequence*. Research Report #47. Washington, DC: NEA.

National Endowment for the Arts. (2009). *Reading on the rise: A new chapter in American literacy*. Washington, DC: NEA.

National Mathematics Advisory Panel. (2008). *Foundations for success: The final report of the National Mathematics Advisory Panel*. U.S. Department of Education: Washington, DC.

National Science Board. (2008). *Science and engineering indicators 2008*. Volumes 1 and 2. Arlington, VA: National Science Foundation.

New Mexico Public Education Department. (2010). *Welcome to New Mexico NAEP!* Retrieved from http://www.ped.state.nm.us/AssessmentAccountability/NAEP/index.html

No Child Left Behind Act of 2001, Pub. L. No. 107-110. (2002). 115 STAT. 1425. Retrieved from http://www2.ed.gov/policy/elsec/leg/esea02/107-110.pdf

OECD (2007). *PISA 2006: Science competencies for tomorrow's world. Volume 1: Analysis.* Paris: OECD Publications.

OECD (2010). *PISA 2009 Results: Learning trends: Changes in student performance since 2000* (Volume V). Paris: OECD Publications. Retrieved from http://dx.doi.org/10.1787/9789264091580-en

Ohanian, S. (2002). *What happened to recess and why are our children struggling in kindergarten?* New York: McGraw-Hill.

Olson, L. (2000, March 29). Study links dropout rate with course requirements. *Education Week*.

Orfield, G. (2006). Forward to J. Lee, *Tracking achievement gaps and assessing the impact of NCLB on the gap*. Cambridge, MA: The Civil Rights Project at Harvard University.

O'Sullivan, C., Lauko, M., Grigg, W., Qian, J., & Zhang, J. (2003). *The nation's report card: Science 2000*. NCES 2003–453. Washington, DC: U.S. Department of Education.

Patte, M. (n.d.) Since when did recess become a dirty word? Examining the impact of NCLB on daily recess time in Pennsylvania elementary schools. Retrieved from http://www.museumofplay.org/taspfiles/ppts/patterecesspaper.ppt

Pellegrino, J., Jones, L., & Mitchell, K. (Eds.) (1999). *Grading the nation's report card: Evaluating NAEP and transforming the assessment of educational progress*. Washington, DC: National Academy Press.

Perie, M., Moran, R., & Lutkus, A. (2005). *NAEP 2004 Trends in academic progress: Three decades of student performance in reading and mathematics*. NCES 2005–464. Washington, DC: Government Printing Office.

Persky, H., Daane, M., & Jin, Y. with Davis, S., Jenkins, F., Liu, H., Qian, J., Thind, S., & Rahman, T. (2003). *The nation's report card: Writing 2002*. Washington, DC: U.S. Department of Education.

Planty, M., Provasnik, S., & Daniel, B. (2007). *High school coursetaking: Findings from the condition of education 2007*. NCES 2007-065. Washington, DC: NCES.

Ramist, L., & Arbeiter, S. (1983). *Profiles, college-bound seniors, 1983*. New York: College Entrance Examination Board.

Rampey, B., Dion, G., & Donahue, P. (2009). *NAEP 2008 trends in academic progress. Reading 1971–2008. Mathematics 1973–2008.* NCES 2009–479. Washington, DC: U.S. Department of Education.

Reese, C., Miller, K., Mazzeo, J., & Dossey, J. (1997). *NAEP 1996 mathematics report card for the nation and the states.* Washington, DC: U.S. Department of Education.

Reese, W. (2002). *Power and the promise of school reform: Grass roots movements during the progressive era*. New York: Teachers College Press.

Roller, C. (n.d.). *No Child Left Behind: A survey of its impact on IRA members*. Newark, DE: International Reading Association. Retrieved from http://www.reading.org/downloads/resources/NCLB_survey_022005.pdf

Ross, E. W. (1996). Diverting democracy: The curriculum standards movement and social studies education. *International Journal of Social Education*, *11*(1), 18–39.

Roy, J., & Mishel, L. (2006). *Rethinking high school graduation rates and trends*. Washington, DC: Economic Policy Institute.

Salahu-Din, D., Persky, H., & Miller, J. (2008). *The nation's report card: Writing 2007*. NCES 2008–468. Washington, DC: U.S. Government Printing Office.

Schmittau, J. (1991). Mathematics education in the 1990s: Can it afford to ignore its historical and philosophical foundations? *Educational Theory*, *41*(2), 121–133.

Schmittau, J. (2004). Vygotskian theory and mathematics education: Resolving the conceptual-procedural dichotomy. *European Journal of Psychology of Education, XIX*(1), 19–43.

Shettle, C., Roey, S., Mordica, J., Perkins, R., Nord, C., Teodorovic, J., Brown, J., Lyons, M., Averett, C., & Kastberg, D. (2007). *The nation's report card: America's high school graduates.* Washington, DC: U.S. Government Printing Office.

Silberman, C. (1970). *Crisis in the classroom*. New York: Random House.

Sizer, T. (1984). *Horace's compromise: The dilemma of the American high school*. Boston: Houghton Mifflin.

Smith, M. (2007). *Leaving NCLB renewal behind*. Retrieved from http://www.educationevolving.org/pdf/mikesmithoped.pdf

Snyder, T., & Hoffman, C. (1991). *Digest of education statistics 1990*. Washington, DC: U.S. Department of Education.

Stedman, L. (1985). A new look at the effective schools literature. *Urban Education*, *20*(3), 295–326.

Stedman, L. (1987). It's time we changed the effective schools formula. *Phi Delta Kappan*, *69*(3), 215–224.

Stedman, L. (1994). Incomplete explanations: The case of U.S. performance in the international assessments of education. *Educational Researcher*, *23*(7), 24–32.

Stedman, L. (1996). Respecting the evidence: The achievement crisis remains real. [Review of *The manufactured crisis*.] *Education Policy Analysis Archives*, *4*(7). Retrieved from http://epaa.asu.edu/ojs/article/viewFile/630/752

Stedman, L. (1997). International achievement differences: An assessment of a new perspective. *Educational Researcher*, *26*(3), 4–15.

Stedman, L. (1998). An assessment of the contemporary debate over U.S. achievement. In D. Ravitch (Ed.), *Brookings papers on education policy: 1998* (pp. 53–121). Washington, DC: The Brookings Institution.

Stedman, L. (2003). U.S. educational achievement in the 20th century: Brilliant success and persistent failure. In R. Weissberg, H. Walberg, M. O'Brien, & C. Bartels-Kuster (Eds.), *Long-term Trends in the Well-Being of Children and Youth* (Chapter 3, pp. 53–76). Washington, DC: Child Welfare League of America Press.

Stedman, L. (2009). *The NAEP long-term trend assessment: A review of its transformation, use, and findings*. Washington, DC: National Assessment Governing Board. Retrieved from http://www.nagb.org/who-we-are/20-anniversary/stedman-long-term-formatted.pdf

Stillwell, R. (2010). *Public school graduates and dropouts from the common core of data: School year 2007–08. First look.* Washington, DC: NCES.

Takahira, S., Gonzales, P., Frase, M., & Salganik, L. (1998). *Pursuing excellence: A study of U.S. twelfth-grade mathematics and science achievement in international context*. NCES 98-049. Washington, DC: U.S. Government Printing Office.

Teachers Network. (2007a). *Results from a survey of teachers on No Child Left Behind*. Retrieved from http://teachersnetwork.org/TNLI/NCLB_Survey_Results.pdf

Teachers Network. (2007b). Survey reveals that only 1% of teachers find No Child Left Behind an effective way to assess the quality of schools. Press release. Retrieved from http://www.teachersnetwork.org/aboutus/pressreleases/nclb_survey.htm

Thompson, B. (2009, January 12). Unexpected twist: Fiction reading is up survey shows reversal of longstanding trend. *Washington Post*, C01.

Toch, T. (1991). *In the name of excellence*. New York: Oxford University Press.

Toch, T., Jerald, C., & Dillon, E. (2008). Surprise—high school reform is working. *Phi Delta Kappan*, *88*(6), 433–437.

U.S. Department of Education. (2003). *The growing numbers of limited English proficient students 1992/93 - 2002/03*. Retrieved from http://www.jillrobbins.com/gwu/227/class6/Growing_LEPpostr.pdf

U.S. Department of Education. (2005, July 14). *Spellings hails new national report card results: Today's news "Proof that No Child Left Behind is working."* Retrieved from http://www.ed.gov/news/pressreleases/2005/07/07142005.html

U.S. Department of Education. (2006). *The growing numbers of limited English proficient students 1995/96 - 2005/06*. Retrieved from http://www.ncela.gwu.edu/files/uploads/4/GrowingLEP_0506.pdf

U.S. Department of Education. (n.d.). *Overview: Four pillars of NCLB*. Retrieved from http://www.ed.gov/nclb/overview/intro/4pillars.html

von Zastrow, C., & Janc, H. (2004). *Academic atrophy: The condition of the liberal arts in America's public schools.* Washington, DC: Council for Basic Education.

Weiss, A., Lutkus, A., Grigg, W., & Niemi, R. with Kulick, E., Swinton, S., & Jerry, L. (2001). *The next generation of citizens: NAEP civics assessments–1988 and 1998*. NCES 2001–452. Washington, DC: U.S. Department of Education.

Weiss, A., Lutkus, A., Hildebrant, B., & Johnson, M. (2002). *The nation's report card: Geography 2001. NCES 2002–484*. Washington, DC: U.S. Department of Education.

Wineburg, S. (1999). Historical thinking and other unnatural acts. *Phi Delta Kappan*, *80*(7), 488–99.

Wood, G. (2004). A view from the field: NCLB's effects on classrooms and schools. In D. Meier & G. Wood (Eds.), *Many children left behind* (Chapter 2, pp. 33–52). Boston: Beacon Press.

Wright, D., Tomlinson, T., & Farris, E. (2005). *Academic requirements and achievement in high schools 1982.* Washington, DC: National Center for Education Statistics.

Zuckerbrod, N. (2007, October 29). More than 1 in 10 high schools in America are "dropout factories." AP wire story. Retrieved from http://www.msnbc.msn.com/id/21531704/

Author

LAWRENCE C. STEDMAN is Associate Professor in the School of Education at the State University of New York at Binghamton. His research interests include historical and contemporary achievement trends; federal policy and school reform; and the transformation of the high school in the 20th century.

Appendix

Minority Achievement Gaps

Characterizing changes in minority achievement gaps is no simple matter as they vary by subject, grade, and age; long-term and main NAEP; and black and Latino students. In reviewing the data, however, I was struck by the fact that most current gaps, especially for older students, are similar to those of 1986–1990 in the long-term NAEP and those of 1990–1994 in the main NAEP (when its trend lines began). In some cases, the gaps had widened during the 1990s, and then closed some, but this means the country is now only back to where it once was.

We need to distinguish gains and gaps. Younger Black and Latino students have improved their reading and math scores, but so have white students, so the gaps have not changed that much since the late 1980s. We could celebrate their score improvement, although they are only now catching up with where white students were 20-40 years ago. Their rate of progress is also much less since NCLB was enacted and the gaps remain large, generally as large as they were fifteen to twenty-five years ago, depending on which comparisons are being made.

The Spellings' Assertion

When the 2004 long-term NAEP results were released, Secretary of Education Spellings heralded the results as "proof that *No Child Left Behind i*s working" and brashly asserted that the gap for younger students had "shrunk to its smallest size in history" (U.S. Department of Education, 2005). As we have seen, however, NCLB is hardly working and the minority achievement gaps remain large, not small. Her claim about historical shrinking was a definite stretch.

In 2004, the black-white math gap for 9-year-olds was a large 23 points, and this was only 2 points smaller than it had been in 1986 (Perie et al., 2005, p. 41). This gap was also statistically comparable to those of the mid-1980s and early 1990s (meaning it fell within a range expected by sampling fluctuations). The Latino-white math gap was a substantial 18 points, only 2 points smaller than it had been in 1982 and statistically matched those of the late 1970s and early 1980s (Perie et al., 2005, p. 43). In other words, the math gaps had not really closed at all.

In reading, the black-white gap for 9-year-olds was still a large 26 points and was only 3 points less than it had been in 1988 (Perie et al., 2005, p. 32). It was, though, substantially smaller than that in several other assessments, so in this case, the data backed Spellings. The Latino-white reading gap, however, was only 3 points smaller than it had been 1984 and was still a substantial 21 points. It was also statistically the same size as those of the late 1980s and early 1990s.

Focusing on younger students distracted attention from what was happening with older ones. The 2004 math and reading gaps for 13- and 17-year-olds *were as large as or larger* than they had been in the mid- to late-1980s (Perie et al., 2005, pp. 33, 35, 42, 44). Instead of proclaiming historically small gaps or that they were closing (Bush, 2008), the Bush administration should have highlighted a large and persistent minority achievement gap.

As it turned out, Secretary Spellings and President Bush were celebrating prematurely. The 2008 long-term NAEP showed that increases by black and Latino 9-year-olds had slowed and that the gaps were essentially unchanged at *all* ages— the changes were tiny (some even

increased) and all were within sampling fluctuations (cf. 1999–2004 and 2004–2008 in Rampey et al., 2009, pp. 14–17, 34–37). The main NAEP showed similar trends, with dramatic slow downs in 4th and 8th grade black and Latino math and reading scores *after* NCLB was enacted (cf. 2000–2003 to 2003–2009 in NCES, 2009, pp. 10, 25; cf. 2000–2002 to 2002–2009 for 4th grade reading and 1998–2002 to 2002–2009 for 8th grade reading in NCES, 2010c, pp. 10, 27).

NCLB and the Closing of the Gaps

Since 2002, minority achievement gaps have barely budged. The evidence comes from subjects assessed by NAEP before and after NCLB was passed and from those it has tracked from 2002 on. U.S. history gaps were essentially unchanged at all grades between 2001 and 2006 (1-3 points; some smaller, but some *larger*) (Lee & Weiss, 2007, p. 14). Science gaps from 2000 to 2005 were smaller at 4th grade (4-5 points), basically the same at 8th (2 points smaller but within sampling fluctuations), and *larger* at 12th grade (3-6 points) (Grigg, Lauko, & Brockway, 2006, pp. 8, 20, 32). Civics gaps between 1998 and 2006 grew smaller at 4th grade for Latino students (by 9 points), but all other gaps were essentially unchanged (1-3 points, some even *larger*) (Lutkus & Weiss, 2007, p. 13). Between 2003 and 2009, math gaps closed by only 1 point at 4th grade and only 3 at 8th (NCES, 2009, pp. 10, 25). In 2009, reading gaps were only 3-4 points smaller at 4th grade and only 1-2 points at 8th (NCES, 2010c, pp. 10, 27). In 2005, the last time 12th graders were tested in reading, the gaps were 1 point larger than they had been in 2000 (Grigg et al., 2007, p. 6). As noted, reading and math gaps in the long-term NAEP had not closed between 2004 and 2008.

U.S. Performance in the International Assessments

In the past decade, three major international programs have reported achievement trends involving U.S. students: Trends in International Mathematics and Science Study (TIMSS) (4th and 8th grades, 1995–2007), Progress in International Reading Literacy Study (PIRLS) (4th grade, 1991–2006), and Program for International Student Assessment (PISA) (15-year-olds in math, science, and reading, 2000–2009). They sample students in both privately- and publically-managed schools.

Mathematics and Science

TIMSS provided some good news. In 2007, U.S. 4th graders performed comparably in mathematics to those in Germany, Denmark, and The Netherlands, while U.S. 8th graders matched those in the Czech Republic, the Russian Federation, Hungary, and England (Gonzales et al., 2008, p. 7). Between 1995 and 2007, U.S. 4th and 8th graders improved modestly in math, gaining 11 and 16 points (p. 8). They did not do as well in science, however. In fact, 4th graders declined initially in science, and while they have recouped a bit, they remain 3 points below where they were in 1995. 8th graders gained 7, less than half what they did in math (p. 33).

The timing is important. 8th graders made most of their math gain over a decade ago (1995–1999) (cf. Gonzales et al., 2008, p. 8 and Mullis, et al., 2004, p. 42). *After* NCLB was enacted, they gained much less in math while their science scores slipped, dropping 7 points between 2003 and 2007 (cf. Gonzales et al., 2008, pp. 33–34 and Martin et al., 2004, p. 44). The weaker position in science may be due to NCLB's lesser emphasis on testing it (only once at 3 school levels as opposed to testing math annually in every grade from 3-8). As schools devoted their energies to meeting external testing mandates in math, they have neglected other areas.

As in NAEP, younger students' math improvement is not that impressive when placed in context. The U.S. 4th and 8th grade gains (11 and 16 points) were small compared to the gains that several other countries made during this time (34 to 57 points) and dwarfed by the large score variations among developed countries at both grade levels—about 130 points (Gonzales et al., 2008, pp. 7–8) (Norway to Hong Kong and Taipei). At the current rate of improvement, a point per year, it will take U.S. 8th graders 60 to 90 years to catch up with students in the top-performing nations (cf. Gonzales et al., 2008, p. 7)—and that is assuming the rate can be sustained and students elsewhere do not improve. The U.S. gains also lose luster when one discovers they were statistically comparable to the marginal 1–4 point gains made by other countries.

The suddenness of the math gains is also an issue. Teaching-to-the-test at the earliest grades likely played a role. All 11 points gained by U.S. 4th graders in math came abruptly between 2003 and 2007—an anomalous jump, similar to that on NAEP, which I discussed in the main text. Skill-drilling, or a recent curricular-test realignment in the international assessments, likely inflated the scores artificially. U.S. gains also seem less impressive when considered in light of the enormous score fluctuations that often show up in the assessments. Many countries displayed erratic performances rather than steady improvement. They had sudden drops—or gains—of about 20–40 points in 4-year spans (Mullis et al., 2004, pp. 42–44). Given that the teaching effectiveness of an entire country is unlikely to have changed so dramatically in a short time, such large and sudden changes in national averages strongly suggest that other factors, such as sampling fluctuations or test unreliability, are at work. Set against them, and dwarfed by them, the U.S. gains of 11–16 points seem unlikely to indicate genuine or substantial math improvement.

At the high school level, U.S. 15-year-olds declined in both science and math during the first half of the 2000s and still lag well behind those in many developed nations (Baldi et al., 2007, pp. 6, 12, 47–48). In science, they dropped 11 points between 2000 and 2006, while in math, they lost 9 points between 2003 and 2006 alone. In the past three years, however, they made sudden and substantial 13-point gains in both subjects (OECD, 2010, pp. 146, 156, 159). Whether these are the seesaw scores that countries often show, a result of curriculum drilling or sampling fluctuations, or the start of sustained, real improvement remains to be seen.

How well U.S. 15-year-olds stack up against those in other nations remains an issue. In science in 2009, they were scoring about the OECD average, but were still 30 to 50 points behind the top-performing countries (OECD, 2010, p. 159). Math performance was worse. U.S. students remained below the OECD average (by 13 points), were more than 25 points behind the median of the 27 countries scoring higher, and more than 50-70 points behind the top performers (OECD, 2010, p. 156). The ranking of the U.S. is not the issue here, but the large achievement gaps are.

Historical Perspective on Math Performance

A long-term perspective is illuminating. In 1982, a year before *A Nation at Risk*, U.S. 8th graders scored around the international average in math (Stedman, 1994, 1997), much as they are scoring now, a generation later. In 1995, the last time 12th graders were tested internationally, U.S. seniors were doing quite poorly in math and science and were among the lowest in the world, just as they had been in the 1960s and 1980s (Takahira et al., 1998, pp. 24, 26, 30–31, 34).

Even some of the best U.S. students in 1995, those taking physics and advanced mathematics, stood low internationally (Takahira et al., 1998, pp. 36, 39, 42, 47, 50). Bronner (1998) reported, "In physics and advanced mathematics, not one of the countries involved—even less well-off nations like Greece, Cyprus and Latvia—scored lower than the United States." AP performance was mixed: U.S. students in AP calculus scored somewhat above the average, while those in AP physics were substantially below it (Takahira et al., 1998, pp. 41, 48). These results contradicted expectations. In the past, the best U.S. students often matched the best in other nations. Bronner (1998) quoted Schmidt, the coordinator of the U.S. assessment,

> This study has burst another myth… Our best students in mathematics and science are simply not world class. Even the very small percentage of students taking advanced placement courses are not among the world's best.

Throughout the 2000s, U.S. 15-year-olds remained substantially behind top-performing nations in both math and science (OECD, 2010, pp. 156, 159). Their weak performance was comparable to those of the international assessments of the early 1980s and mid-1990s (Stedman, 1994, 1997).

Reading

In the past, U.S. students typically did better internationally in reading than in math and science, but their performance declined in the 1990s and early 2000s. 4th graders' reading scores fell a substantial 10 points in the 1990s and slipped a bit more between 2001 and 2006 (Martin et al., 2003, p. 15; Mullis et al., 2007, p. 44). In 1991, U.S. 15-year-olds were substantially *above* the international scale average in reading (35 points higher) and among the best countries in the world, ranking 9th (Elley, 1992, p. 24). By 2003, they had dropped *below* the international scale average and fallen behind 17 nations (Lemke et al., 2004, p. 92). Between 2000 and 2003 alone, they lost almost 10 points. By 2009, they had recovered some of the lost ground, with a 5-point improvement (OECD, 2010, p. 146).

In all three areas—reading, math, and science—the performance and relative standing of U.S. students internationally slips as they go through school (cf. Gonzales et al., 2008, pp. 7, 32 and Baldi et al., 2007, pp. 6, 12; cf. Mullis et al., 2007, p. 37 and Lemke et al., 2004, p. 92).

Judging the International Findings

My discussion here has centered on the declines in U.S. performance and the large size of the achievement gaps between it and other countries, *not* on its national ranking (first, second, third, next-to-last, last, etc.). It is important to distinguish achievement gaps and rankings. National rankings are a problematic way of judging a country's performance (Stedman, 1997). Rankings can differ greatly, while achievement may be similar. Focusing on the achievement gaps—the actual differences in performance—avoids those problems. The gaps have often been large and educationally significant. Scoring *well below* the international averages or *well behind* the top-performing countries are serious matters, no matter where the U.S. ranks.

My primary concerns, however, are not the gaps, but that U.S. students lost ground internationally during much of the standards era and have repeatedly demonstrated weak knowledge and skills in the international assessments. The reason I described the relative performance and low standing of U.S. students is because those have been an ongoing concern of many observers, have played a central part in national educational debates, and show how far

away the country is from achieving the Goals 2000 goal of having U.S. students be first in the world in math and science.

Let me be clear. I am certainly not embracing the Goals 2000 call for U.S. educational supremacy. While the gaps between the U.S. and top-performing countries are a serious issue, we should still reject the Goals 2000 call for U.S. students to be first in the world in math and science achievement. That grew out of a misguided Cold War-type nationalism and a push to maintain U.S. economic hegemony. It was even an unnecessary goal because another goal called for all students to master challenging material in the major subjects, including math and science. We also should reject the view that economic globalization requires U.S. students be among the world's best, so that U.S. companies have the human capital needed to compete in the new world order. Such arguments are simplistic and unsubstantiated. Still, U.S. students should be doing better in absolute terms in the international assessments (as measured by their knowledge and understanding not their ranking), if only for the sake of their intellectual and personal development, functional literacy skills, and basic academic preparation.

Center on Education Policy Findings on NCLB

As noted in the main text, the Center on Education Policy (CEP) conducted a multi-year series of nationally representative studies to investigate the impact of NCLB. CEP found that school districts have changed their curricula to align with NCLB-mandated tests (McMurrer, 2007, p. 2). In its year-4, post-NCLB survey, for example, 71% of districts had increased reading and math time, while reducing it in another area (CEP, 2006, p. vii). In its year-5 survey, CEP again found a large majority of districts (62%) had increased time on reading and/or math, while almost half of all the districts surveyed reported increasing time in reading and/or math *and* reducing it in other areas; on average, by about a third (McMurrer, 2008, pp. 2, 4). Over a third of all districts had decreased time in social studies, about a fourth in science, a fifth in recess, and a sixth in art and music. The weekly decreases in each area were about an hour or more.

The central finding about reduced instructional time has been misreported. Some of the blame for this falls on CEP's shoulders as its press release about the study was subtitled: "Instructional time for subjects not tested under No Child Left Behind has fallen by nearly one-third since law was passed" (CEP, 2007). But this was true *only* for the districts that had *increased* time in reading or math (or both)—not *all* districts. CEP did *not* present the overall changes in average time for *all* districts and the data in its charts were contradictorily labeled. In Table 1 in McMurrer (2008, p 2), for example, the average decreases in time are labeled as applying to *all* districts that decreased time, whereas in Table 3, the same data is presented as applying *only* to those districts that had also *increased* time on "ELA and/or math" (p. 4). An important limit of the CEP studies is that these are district, not school or teacher, reports and so the time estimates are likely not that accurate a measure of what is transpiring in schools.

Public High School Graduation Rates

Determining high school graduation rates is a complex and contentious matter. The rates—and trends—vary greatly depending upon the assumptions, methods, and data sets used. Using a different approach than Heckman & LaFontaine (2007), Roy and Mishel (2006, p. 24) found an overall rate of over 80% and around 70%, not 50%, for minority students. Heckman and LaFontaine are critical of their analysis and findings. (The graduation rates discussed here are for students receiving *regular public* school diplomas, not GEDs or private school ones.)

The decline in the graduation rate given in the main text—from about 80% to 75% between 1984 and 2001—reflects Heckman and LaFontaine's finding (see their Figure VII) as well as a calculation using their method of comparing the number of *8th graders* in public schools in the fall of a year to the number of public school graduates 5 years later (a rationale for this method is given in their article). Thus, 8th graders in the fall of 1979 are compared to 1984 graduates— and 1996 8th graders to 2001 graduates. (I calculated the rate using data from NCES, 1993, p. 53; NCES, 2010a, pp. 71, 164. It declined from 78.7% to 75.5%. My calculation does not take into account 8th graders in *ungraded* elementary schools or those who left for private schools, private school 8th graders who later graduated from public school, or changes in retention rates over time.)

The old method of determining graduation rates, which is still widely used, is problematic. It compares the number of 9th graders to the number of graduates 4 years later. The issue is that a large 9th grade retention rate inflates the base and, thus, *underestimates* graduation rates, especially for minority students. The new 8th grade method, however, *inflates* the rates as it now includes graduates from more cohorts. So, the true graduation rates may be even lower.

To solve such problems, NCES now uses what they call the "averaged freshman" graduation rate, which averages the number of 8th, 9th, and 10th graders over three consecutive years to determine the base (Stillwell, 2010, pp. 1 and 6, note 1). This method does show a *lower* overall graduation rate. It also confirms the decline in graduation rate during the standards era—from about 75% in 1984 to 72% in 2001 (NCES, 2010a, p. 164). Since then, this rate has risen and leveled off at about 75% but, as noted in the text, increases post-NCLB passage are suspect.

Heckman and LaFontaine's (2007, p. 17) description that the graduation rate "steadily declined" from the early 1970s is an overstatement—it increased for a few years leading up to *A Nation at Risk* (Heckman & LaFontaine, 2007, Figure VII).

A major success of *A Nation at Risk*—the increase in state graduation requirements—has been linked to increases in the dropout rate. In one study, Lillard and DeCicca (2001, p. 459) found that, net of other factors, students "are more likely to drop out of high school in states with higher course graduation requirements." In another, involving analyses of several state and cohort data sets, Lillard and DeCicca (2001) determined that increased state graduation requirements had, *conservatively*, increased dropouts 3% to 7.4%. Another estimate was 4-8%. With a dropout rate of about 20% in 1984, an 8% impact would be about a 1.6 percentage point increase in dropouts. Given that the graduation rate declined about 3.2 percentage points from 1984 to 2001, this means that increased graduation requirements may have caused *half* the increase in the dropout rate.

One federal indicator has produced contradictory results, showing a low dropout rate and a decline among 16- to 24-year-olds (Cataldi, Laird, KewalRamani, & Chapman, 2009, p. 33). This "status dropout rate," however, is misnamed, as it does not count GED recipients as dropouts. It also depends upon self- or household-reporting in census surveys, not school records, and by covering teenagers, includes some who will be returning to high school. Most other methods of determining graduation rates converge and show that dropout rates are large and increased during much of the standards era (see Cataldi et al., 2009, pp. 1–2; Kaufman, Alt, & Chapman, 2004, p. 4; NCES, 2010a, p. 164).

The Proficiency of High School Students

The standards movement has fallen well short of its goal of having 100% of the nation's students be proficient. As documented in the main text, although both Goals 2000 and NCLB called for all students to be proficient (NCLB, 2002; NEGP, 1995, p. 36), only about a fourth of high school students have done so in the main NAEP. The long-term NAEP uses a different set of proficiency levels, defined as reaching 300 and 350 scale scores. On average, less than half the nation's high school students have reached the 300 level and very few—*only 6%*—reach the 350 level, the levels purportedly needed for business, higher education, and government (Campbell et al. 2000, pp. 21, 23, 25; Rampey et al., 2009, pp. 13, 33; Stedman, 1998, 2009). The predictive validity of these levels is undetermined, but fueling the concern over student performance is that some of the items defining the levels are basic (Stedman, 1998, p. 77). The 350 level in mathematics, for example, includes "routine problems involving fractions and percents" yet, over the years, only 5-8% of high school students have reached it (Rampey et al., 2009, pp. 32–33).

The standards movement also did little to improve high school proficiency. Since the early 1990s, the percentages of proficient 12th graders in the main NAEP have risen only 1-2 points in U.S. history, civics, and writing (Lee & Weiss, 2007, p. 9; Lutkus & Weiss, 2007, p. 9; Salahu-Din et al., 2008, p. 37). In math, the percentage jumped at the start of the 1990s, then rose only 2 points between 1992 and 2000 (Braswell et al., 2001, p. 26). On the new math assessment, though, the percentage jumped 3 points between 2005 and 2009 (NCES, 2010b, p. 26). In other areas—reading, science, and geography—the percentages have even dropped several points and are now *less* than they were in the early- and mid-1990s (NCES, 2010b, p. 9; Grigg, Lauko, & Brockway, 2006, p. 31; Weiss et al., 2002, p. 21).

Similarly, in the long-term NAEP, the percentages of 17-year-olds reaching the 300 and 350 levels have not changed much (Campbell et al. 1998, pp. viii, xii; Campbell et al., 2000, p. 25; Campbell, Reese, O'Sullivan, & Dossey, 1996, p. 189; Rampey et al., 2009, pp. 13, 33). While there were some improvements in the 1980s at the 300 level, trends have been flat or even declining since then, especially in literacy. The percentages achieving 350 in reading and math have been roughly constant for three decades (Rampey et al., 2009, pp. 13, 33).

NAEP Item Review: Student Performance by Subject

As noted in the main text, evaluators have questioned the validity of the NAEP levels. Analyzing performance on individual test items, however, bypasses these issues and shows that deep problems persist in all the major subjects. In this section, I review performance on items from recent NAEP assessments. Except where noted, the results come from NAEP's online database (NAEP, 2010c). Data are for both public and private school students. For earlier findings, see Stedman (1996, 1998, 2003). I concentrate on high school seniors and 17-year-olds because their performance represents the culmination of twelve years of U.S. schooling.

Student performance certainly has good aspects. In each tested area, students do well on a number of items. High school students are familiar with several major historical documents and countries. Most answer correctly civics items pertaining to rights, responsibilities, and the law (Weiss et al., 2001, p. 69). They can add, subtract, and read some simple graphs. Most spell and punctuate correctly.

We should not judge their performance too harshly as many of us gained some of this

knowledge after high school, in college courses, by reading newspapers and historical biographies, travelling, following current events, etc. Some students also might score higher on NAEP and the international tests if the stakes were higher and they were more motivated. Still, such low-stakes tests reduce test anxiety and the stereotyping threats that harm performance. Unlike the long and arduous SATs, NAEP's tests are short ones, with much middle-school material, making it easier for students to do well. Overall, high school seniors struggle with much material that is basic and that they should have been learned during their 12 years of schooling (Stedman, 1998).

U.S. History Performance

High school students continue to lack basic historical information. Most 12th graders do not know why FDR tried to pack the Supreme Court or what the Lincoln-Douglas debates were about. Over 85% cannot give a reason the U.S. was in the Korean War and why it mattered (Lee & Weiss, 2007, p. 28). More than half do not know what Social Darwinism is or recognize Nat Turner or other leaders of slave revolts. These problems mirror those in previous NAEP assessments. In the mid-1990s, most 17-year-olds did not know which countries were involved in the Camp David accords and more than half did not realize that containing communism dominated U.S. foreign policy after World War II (Stedman, 1996).

Geography Performance

High school seniors still have limited geography knowledge and weak map skills. While they do well on multiple-choice items pertaining to earthquakes, over 80% cannot use a set of maps to identify where most people live in Australia and give a reason for it (NAEP, 2010c). Three telling examples of their performance come from the latest NAEP geography report card (Weiss et al., 2002, pp. 101–106). Nearly two-fifths cannot identify Hinduism as the leading religion in India. More than half cannot give two reasons the Tigris-Euphrates region was the site of early civilization. Over 80% cannot give a reason why two population pyramids differ. As noted in the main text, despite the Vietnam War and recurring Mideast crises, almost two-thirds in past assessments could not identify Southeast Asia, the Persian Gulf, or Saudi Arabia. For more information on performance in past geography assessments, see Stedman (1996, 1998).

Civics Performance

Civics knowledge has been weak for several decades. The 1998 assessment showed that most 12th graders did not know the meaning of referendum, bicameralism, or PAC (56%, 62%, 66%) (Weiss et al., 2001, pp. 14, 70). Most did not know that treaties must be ratified by the Senate (61%) or that federal judges are nominated by the President and confirmed by the Senate (58%) (p. 70). The 2006 assessment revealed that half the seniors did not realize that Congress shares foreign policy power with the President (NAEP, 2010c). Over three-fourths could not explain a 1960s political cartoon depicting the domino theory (Lutkus & Weiss, 2007, p. 29; NAEP, 2010c). Even about half of the "proficient" students had trouble with this item. Students have surprising gaps in their knowledge of constitutional issues and race relations. In 1998, only 40% knew that states, not federal laws or regulations, had segregated schools (NAEP, 2010c). In 2006, only about half knew that *Brown v. Board of Education* had led to a struggle over states' rights (55%) (NAEP, 2010c).

Science Performance

Seniors also have striking weaknesses in earth science, chemistry, and biology (NAEP, 2010c). Over three-fourths, for example, do not know that nitrogen and oxygen are the primary gases in the Earth's atmosphere. Three-fourths cannot balance a basic chemical equation. Over half cannot explain how the body regulates temperature. *Over 90% could not explain fully why menstruation ceases during pregnancy.* (Girls were more likely to give *partially* correct responses, but about three-fourths of both genders gave unsatisfactory or incorrect answers.) Over half did not know the Doppler effect involves wavelength changes. In 2000, three-fifths could not pick the reason the Sun appears slightly larger in January than July (O'Sullivan et al., 2003, p. 27).

Math Performance

High school students still struggle with middle-school math, including simple algebra, routine word problems, and several types of basic problems involving fractions and areas. In 2004, only about a third of 17-year-olds, for example, could convert the decimal .029 to a fraction or find the area of an L-shaped rectangular region (NAEP, 2010c; Perie et al., 2005, p. 88). *Less than half could find the product of 3 and 2 1/3.* In 2005, less than a third of seniors could find where two lines intersect and only about a fourth could determine the weighted average of two groups. In 2008, less than 60% of 17-year-olds could find the area of a square given its perimeter and less than a third could find the unit profit from a candy bar sale.

Even though their scores have improved, younger students, including many labeled "proficient," still have trouble with basic math. In 2005, more than two-thirds of 8th graders could not determine what percentage of a bill a tip was. About three-fourths do not know a rectangle is a parallelogram. In 2007, two-thirds could not convert a Fahrenheit temperature to Celsius even when given the steps and a calculator. In 2009, 60% could not find the hypotenuse of a right triangle. 70% could not determine the next number in a sequence where each was multiplied by 8. 84% could not determine how many square tiles were needed to cover a rectangle.

U.S. 8th graders also had problems with basic math and algebra in the international assessments. For example, while 81% can match a pie chart to a shaded rectangle, only 38% can determine how much money is needed to buy a pen and 2 pencils and only 40% can make a bar chart when given a pie chart and the total (Gonzales et al., 2008, pp. B-5–B-8).

Over twenty years ago, math educators who examined NAEP results determined that students—at all ages—had "superficial learning" and "major deficiencies" (Carpenter et al., 1988, pp. 40–41). Alas, the same is true today.

Writing Performance

Writing assessments have provided some of the most direct evidence of problems because NAEP evaluates papers holistically using a 6-point scale ranging from excellent to unsatisfactory (Persky et al. 2003, pp. 92–94; Salahu-Din et al., 2008, p. 5). Students are graded on narrative, informative, and persuasive writing tasks.

Over the past decade, only about a fourth of the papers written by high school seniors have been judged skillful or excellent (Greenwald et al., 1999, p. 132; NAEP, 2010c; Persky et

al. 2003, pp. 139, 147, 155; Salahu-Din et al., 2008, p. 44). More of them, about a third, have been considered inadequate, being rated uneven, insufficient, or unsatisfactory. Even the so-called "sufficient" papers had problems. They provided only "some pertinent details" or "some support" for their statements or stories, frequently lacked transitions or had none at all, often had "simple and unvaried" word choice and sentence structure, and inaccurately used words (Persky et al. 2003, pp. 92–94; Salahu-Din et al., 2008, p. 45). They also had errors in grammar, spelling, and punctuation, although those "did not interfere with understanding" (Persky et al. 2003, pp. 92–94). Overall, NAEP finds that less than a fourth of 12th graders are proficient writers (Salahu-Din et al., 2008, p. 37).

These weak results are similar to those in the past (Stedman, 1998, 2003). They are surprising because the writing tasks are simple, inviting ones; require only short, single paragraph or single page responses; and are treated as first drafts (Salahu-Din et al., 2008, pp. 5, 42). Students are even given a short, clear writing guide to help them do the tasks.

In spite of a generation of standards efforts, therefore, students still struggle in diverse subjects and skill areas. The standards movement has failed in its central mission to produce high levels of academic achievement and knowledgeable, proficient students.

Critical Education

criticaleducation.org

ISSN 1920-4175