



# Cato Handbook for Policymakers

**CATO**  
INSTITUTE

7TH EDITION

## 20. K-12 Education

### Congress should

- recognize that its education programs have proved incapable of achieving their stated goals;
- understand that its failed education programs have cost American taxpayers \$1.85 trillion since 1965; and
- phase out all its K-12 education programs over three years, resulting in a \$70 billion per year tax cut for the American people.

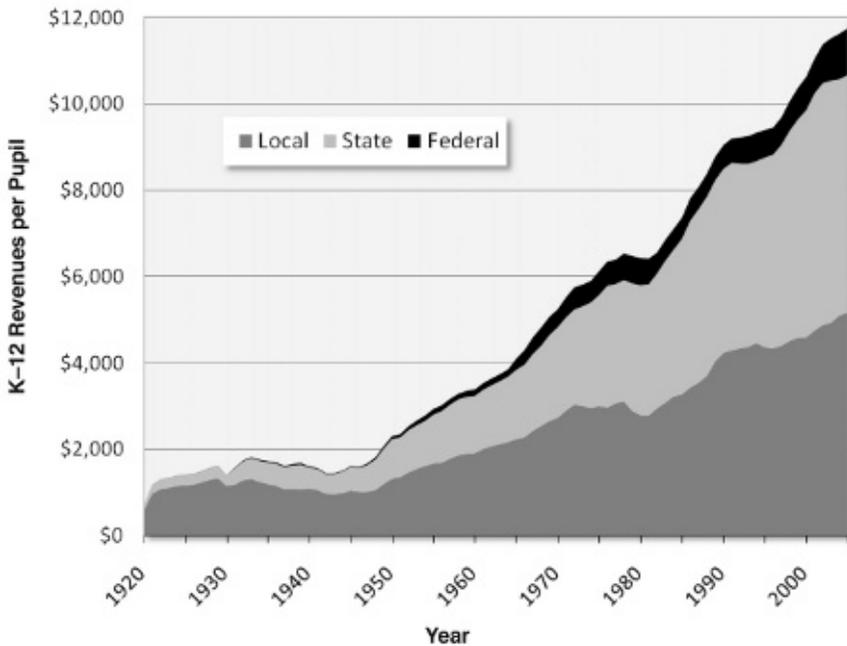
Ultimately, the proper federal role in education is whichever one would best serve the public in the long run. That does not mean Congress should ignore the Constitution, which plainly leaves authority over education in the hands of the states and the people. What it means is that if a careful review of the evidence demands a federal role that is not currently authorized, the Constitution could and should be amended to allow it. Conversely, if federal education programs have consistently failed to produce results commensurate to their cost, they should be discontinued.

But what does the evidence show? This chapter reviews federal involvement in education to date, assessing the outcomes of federal programs in light of the considerable resources taxpayers have invested in them. It concludes by offering the simple policy recommendations that follow inexorably from that review.

### Have Federal Education Programs Worked?

As shown in Figure 20.1, federal revenues constituted less than 1 percent of public school budgets before the 1930s, and grew only imperceptibly until the post-World War II period. That negligible level of spending was matched by a complete lack of federal interest in shaping what went on in America's classrooms. It wasn't until the National Defense Education

**Figure 20.1**  
**K–12 Revenues per Pupil, by Level of Government, 1920 to 2005**



SOURCES: Thomas Snyder, Sally Dillow, and Charlene Hoffman, *Digest of Education Statistics 2007* (Washington: U.S. Department of Education, 2008), Table 162. Missing values are linearly interpolated. Historical consumer price index inflation factors are from <http://oregonstate.edu/cla/polisci/faculty-research/sahr/sahr.htm>.

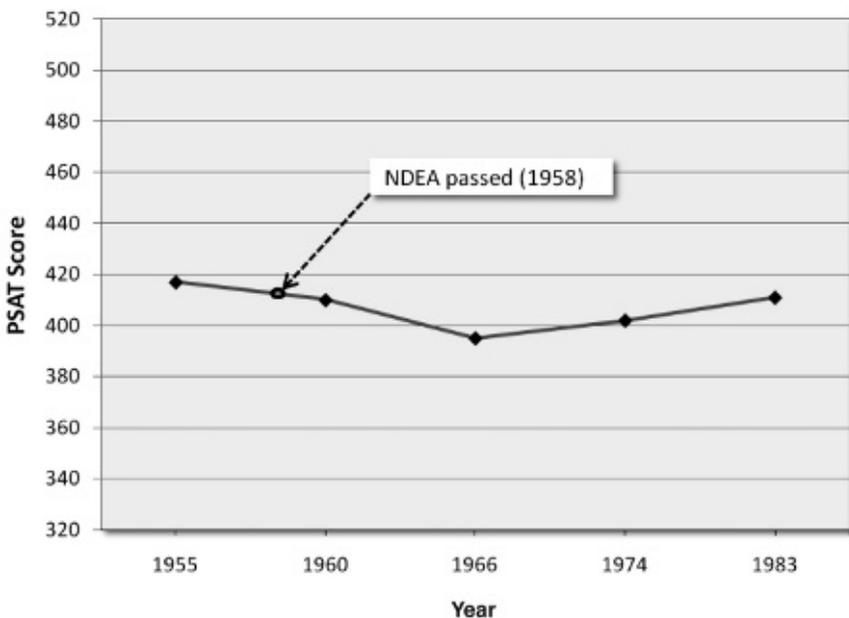
Act of 1958, aimed at beefing up math and science instruction after Russia’s launch of the satellite Sputnik, that Congress sought to shape public school instruction.

What did the NDEA do, and did it work? Some NDEA funding purchased new equipment for science classrooms, but much of it went toward the development of new math and science curricula by experts in colleges of education, new teacher training programs, and the hiring of experts in these fields as curriculum supervisors within the nation’s public school systems. What effect did these expenditures have? That question was addressed in 2006 by the Science and Technology Policy Institute under a federal research contract. According to the STPI study, public school district officials felt that NDEA-funded programs had done some good but that states “found it impossible to identify direct causal links among the myriad of variables influencing public education”—bureaucratese for “who knows?”

Of course the most obvious and meaningful test of the NDEA's effectiveness is to look at academic achievement in math and science before and after the law's passage—something that the STPI study failed to do. Fortunately, nationally representative data for the period are available from the Educational Testing Service, which administered its Preliminary Scholastic Aptitude Test to nationally representative samples of high school juniors periodically between 1955 and 1983. The mathematics results of these little-known "national norm studies" appear in Figure 20.2 (note that the PSAT does not include a science portion).

The results depicted in Figure 20.2 indicate that mathematics performance declined slightly during the latter half of the 1950s, and that this decline actually accelerated from 1960 to 1966, after the NDEA was passed. The avowed belief of school district officials that the law was working is not supported by these results.

**Figure 20.2**  
**Math Scores, National Norm PSAT Studies (11th graders),**  
**1955–83**



SOURCE: Charles Murray and Richard J. Herrnstein, "What's Really behind the SAT-Score Decline," *Public Interest*, no. 106 (Winter 1992): 32–56.

Looking at the apparent turnaround beginning in 1966, it might be theorized that the new programs and staff put in place by the NDEA did not begin to work until a decade or so after the law was passed. Perhaps, after finally getting their bearings, they contributed to a lasting upward trend in mathematics performance. Whether or not that theory is plausible in principle, it, too, is contradicted by the evidence.

Science and mathematics scores on the Long-Term Trends tests of the National Assessment of Educational Progress became available at the end of the 1960s and beginning of the 1970s. These tests, administered at regular intervals to nationally representative samples of 9-, 13-, and 17-year-old students, offer the best picture of changes in average student achievement over time. The results for younger students are useful for illustrating changes in the relative achievement at different levels of public schooling—how the performance of elementary, middle, and high school children varied over time. But since the goal of the NDEA was to better prepare students for success in challenging college work after high school, it is to the performance of 17-year-olds that we must turn to assess the law's effectiveness.

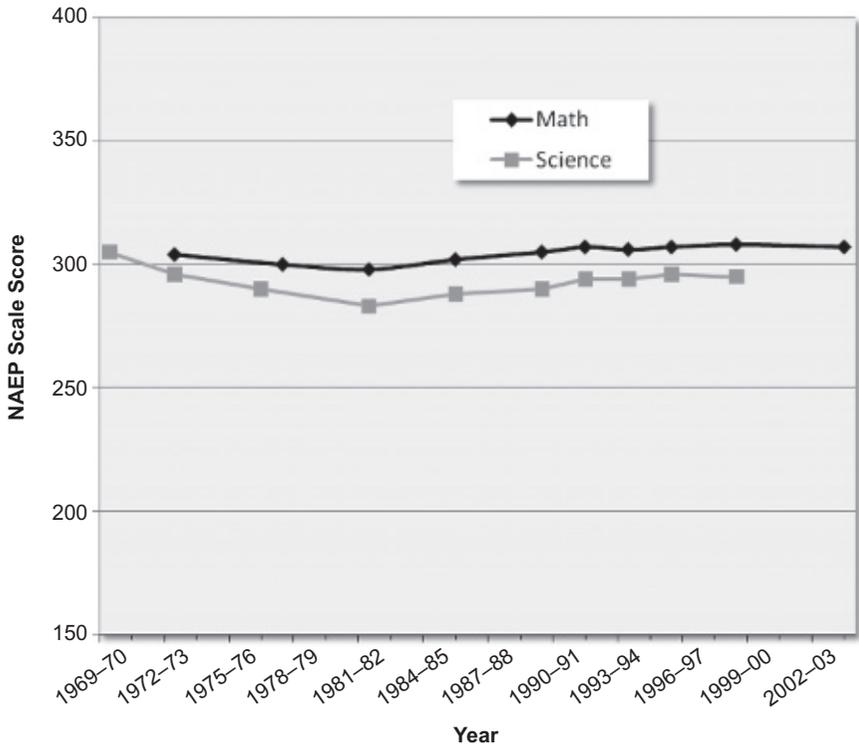
As the results shown in Figure 20.3 illustrate, the rising math trend on the PSAT through 1983 is not found on the NAEP mathematics test. On the contrary, NAEP math scores declined from the early 1970s through the early 1980s. Overall, math scores are statistically unchanged over the past three decades. In science, a striking decline occurred through the early 1980s from which scores never fully recovered. At the end of high school, student performance is statistically significantly worse in science today than it was when the NAEP test was first administered in 1969–70.

Despite the NDEA's failure to improve mathematics and science achievement, Congress and the president decided to repeat the same ineffective strategy nearly half a century later, with the America COMPETES Act of 2007 (an acronym for Creating Opportunities to Meaningfully Promote Excellence in Technology, Education, and Science). It is not clear why legislators believed the results would be different this time around. It is not even clear that legislators were aware of the earlier failure of the NDEA.

## **Head Start and ESEA**

In 1965, the federal government created two new programs aimed at closing the achievement gaps between high- and low-income children and between white and black children: the Elementary and Secondary

**Figure 20.3**  
**NAEP Long-Term Trends Results (17-year-olds), 1969-70 to 2003-04**



SOURCES: Rebecca Moran and Anthony D. Lutkus, *NAEP 2004 Trends in Academic Progress: Three Decades of Student Performance in Reading and Mathematics* (Washington: U.S. Department of Education, 2005), p. 17; Jay R. Campbell, Catherine M. Hombro, and John Mazzeo, *NAEP 1999 Trends in Academic Progress: Three Decades of Student Performance* (Washington: U.S. Department of Education, 2000), p. 9.

Education Act and a preschool program known as Head Start. As its name implies, Head Start was intended to prepare disadvantaged students to perform better academically once they entered the K-12 system. According to the most recent large-scale study of the program, authored by Michael Puma and others, Head Start has some small positive effects in the early language skills of 3-year-olds, but no effect on their mathematics skills, when compared with a control group of students not receiving Head Start services. Among 4-year-olds, this same federally funded study found fewer language skill improvements and still no mathematics improvements.

These findings are consistent with a comprehensive literature review of the Head Start research released by the Department of Health and

Human Services in 1985. That review described some initial academic gains among preschoolers while they were participating in the program, but a gradual atrophy of such gains as students progressed through the elementary and higher grades. In fact, according to the 1985 study, there were no lasting academic achievement benefits to Head Start:

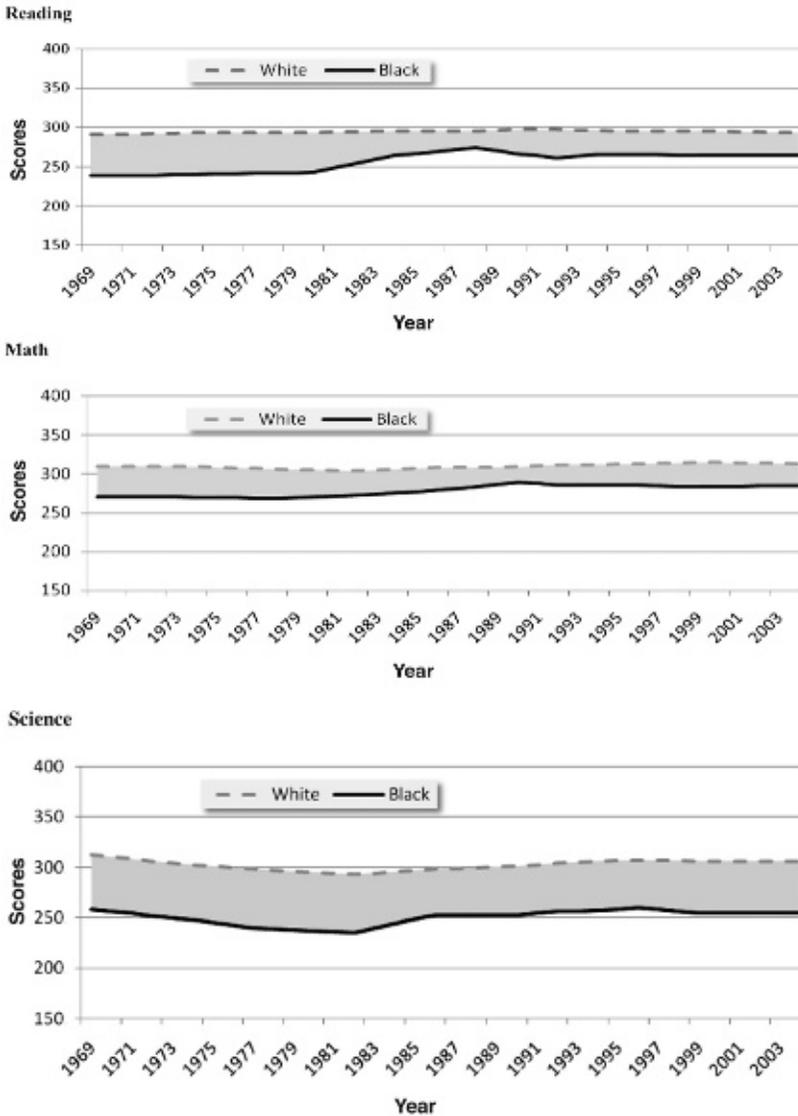
Children enrolled in Head Start enjoy significant, immediate gains in cognitive test scores, socioemotional test scores, and health status. In the long run, however, cognitive and socioemotional test scores of former Head Start students do not remain superior to those of disadvantaged children who did not attend Head Start.

Although the 1985 DHHS study reviewed over 200 research papers, a few, both before and since, have deviated from the overall pattern of vanishing benefits over time. Even these exceptional findings, however, have considerable caveats. A study by three UCLA economists published in 2000 (“Longer Term Effects of Head Start”) reported that the children of white high school dropouts who participated in Head Start had higher earnings between the ages of 23 and 25 than their siblings who did not participate in the program. No such earnings effect was found for African Americans or for other age ranges. The same study, did, however, find higher graduation rates among whites and among African-American males who had participated in the program (again, compared with siblings who had not).

Head Start represents a relatively small fraction of total federal spending on K–12 education. The bulk of that spending is made under the auspices of the Elementary and Secondary Education Act, which comprises hundreds of different programs. It directs funding to states and school districts for everything from early literacy to dropout prevention, and from “comprehensive school reform” to teacher training. All these programs are collectively intended to advance the law’s aim of closing the racial and economic achievement gaps. After more than 40 years of programs providing services from preschool to high school, the most meaningful way to determine the effectiveness of the federal government’s interventions is to look at how the achievement gaps have changed over time. Once again, it makes the most sense to look at students at the end of high school to see the cumulative effects of all federal programs (if any).

Figure 20.4 charts scores for black and white 17-year-olds on the three main NAEP tests (reading, mathematics, and science), to get an idea of changes in the racial achievement gap over time. As is evident from these graphs, the overall gap remained essentially unchanged through 1980.

**Figure 20.4**  
**NAEP Long-Term Trends Scores for 17-Year-Olds, by Race,**  
**1969-2003**



SOURCES: Rebecca Moran and Anthony D. Lutkus, *NAEP 2004 Trends in Academic Progress: Three Decades of Student Performance in Reading and Mathematics* (Washington: U.S. Department of Education, 2005), pp. 33, 42; Jay R. Campbell, Catherine M. Hombro, and John Mazzeo, *NAEP 1999 Trends in Academic Progress: Three Decades of Student Performance* (Washington: U.S. Department of Education, 2000), p. 37.

That is 15 years after the passage of both the ESEA and Head Start, time enough for the children being tested to have passed through all these federal programs from preschool through the end of high school. Nevertheless, the gap had not narrowed.

Then, across subjects, the racial achievement gap among 17-year-olds shrank in the early to mid-1980s. That trend subsequently ceased, and even reversed itself in the 1990s, though the reversal was not uniform in timing or in magnitude across subjects. In reading and mathematics, the achievement gap remains smaller than it was when the tests were first administered. In science, there is no longer a statistically significant difference in the gap between the earliest and most recent test results.

If federal education programs were having their intended effect, we would expect the gaps to have narrowed more or less uniformly over the past 40 years. That has not happened. To credit federal programs with the narrowing of the early 1980s, or to blame them for the subsequent widening, it would be necessary to point to major changes in federal programs in the years leading up to those fluctuations. There are no obvious federal policy shifts that one might point to that could explain these changes. The gap fluctuations must therefore be explained by other factors, whether social, economic, or relating to state-level policy.

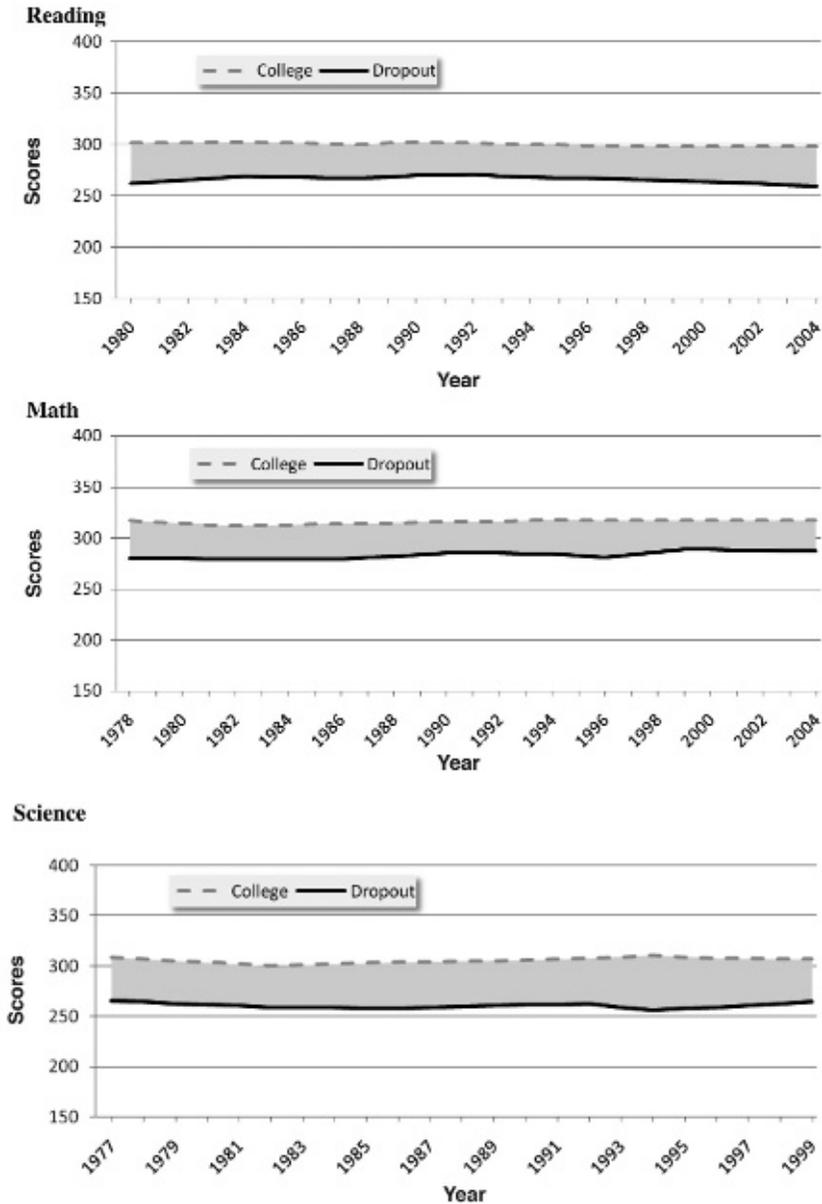
What about the gap in performance between rich and poor Americans, the reduction of which was the principal stated aim of both Head Start and the ESEA? Though breakdowns of NAEP scores by family income are unavailable, we can get a rough idea of them by looking at scores according to parental level of education (income and education level being strongly correlated). Figure 20.5 presents the score gaps between students whose parents graduated from college and those whose parents did not complete high school.

As Figure 20.5 makes clear, the gap in achievement between children of college graduates (generally wealthier) and those of high school dropouts (generally poorer) has not narrowed in either reading or science. The gap reduction that has occurred in mathematics amounts to barely 1 percent of the 500-point scale, and fluctuations in that gap are not obviously correlated with any particular shifts in federal policy, making it doubtful that federal policy played a decisive role even in that tiny change.

## **No Child Left Behind**

Disappointed by the ineffectiveness of federal programs despite generations of effort and a massive outlay of taxpayer dollars, Congress and

**Figure 20.5**  
**NAEP Long-Term Trends Scores for 17-Year-Olds, by Parents' Level of Education, 1977-2004**



SOURCES: Rebecca Moran and Anthony D. Lutkus, *NAEP 2004 Trends in Academic Progress: Three Decades of Student Performance in Reading and Mathematics* (Washington: U.S. Department of Education, 2005), pp. 37, 46; Jay R. Campbell, Catherine M. Hombo, and John Mazzeo, *NAEP 1999 Trends in Academic Progress: Three Decades of Student Performance* (Washington: U.S. Department of Education, 2000), pp. 50-51.

President Bush decided to shift the emphasis of ESEA programs in 2002, reauthorizing the law under the name No Child Left Behind. In this reconstituted form, states were encouraged to test students in reading and mathematics and to focus their improvement efforts on those schools failing to show adequate progress on those tests from year to year. Has NCLB proved more effective than past federal initiatives?

The National Assessment of Educational Progress has been the one source of nationally representative evidence that has been widely used to assess NCLB's effectiveness. "High-stakes" state-level results are also available, from state tests that are used to judge the performance of schools and school officials, as well as students, but they are not nationally representative and are widely regarded as unreliable measures of changes in student performance over time. Their unreliability is due to reductions in test difficulty over time; narrow teaching to the specific test content; and on occasion, even outright cheating by school officials.

As already discussed, it would be most meaningful to look at trends for students in the final years of public schooling, to observe NCLB's lasting effects (if any). But older students have thus far had little exposure to key NCLB provisions, which are targeted chiefly at the lower grades. Hence, data for elementary and middle school students must be examined for the time being.

According to the NAEP's Nation's Report Card series of tests, 4th-grade reading scores increased by less than one-half of 1 percent of the 500-point scale between 2002 and 2007. As it happens, 4th-grade scores had increased by just over 1 percent between 2000 and 2002, before NCLB was passed. So if NCLB has had an effect at all, it has been to slow the rate of improvement that existed immediately before its passage. The same, incidentally, goes for the black–white score gap, which shrank by 0.8 percent of the score scale between 2000 and 2002, but only by 0.6 percent of the scale in the five years after NCLB's passage. At the 8th grade, students' reading scores fell by one point (0.2 percent of the score scale) between 2002 and 2007, and the black–white score gap remained unchanged.

In mathematics, the story is much the same. In the 4th grade, scores improved by 1 percent between 2003 and 2007, but had improved by nearly 2 percent between 2000 and 2003, before the NCLB could have had any effect. The racial achievement gap shrank by 0.6 percent after NCLB, but by 0.8 percent from 2000 to 2003. Scores for 8th graders rose by 0.4 percent after NCLB, but had improved by 1 percent between 2000

and 2003. The story is once again the same with respect to the achievement gap between black and white students.

It is worth noting that gains in NAEP scores at the 4th- and 8th-grade levels have historically failed to persist through the end of high school, as is evident in the flat trend lines for 17-year-olds shown in the figures earlier. This implies that no additional learning is taking place over the course of children's public schooling, but that some learning has merely shifted to lower grades.

The tiny and likely evanescent post-NCLB gains for younger children, which are smaller than those that occurred immediately before NCLB's passage, represent the good news, such as it is. There is also bad news, which NCLB advocates and the media have uniformly ignored. Two sets of international tests, the Program on International Student Assessment and the Program on International Reading Literacy Survey, offer nationally representative data on trends in U.S. student achievement over the NCLB's lifetime. In every subject, on both tests, the scores of American students have declined.

PISA was first administered to 15-year-olds in 2000, testing them on mathematics, reading, and science. Students in the United States initially earned an overall math score of 493 on the 1,000-point scale, 7 points below average, placing us 18th out of the 27 participating countries. Three years later, PISA results showed no significant change in U.S. math performance. But according to the latest report, the United States suffered a significant decline in mathematics achievement between 2003 and 2006, a period in which an NCLB effect could reasonably have been expected. We now score 474—in 25th place among the 30 participating countries.

PISA also tests students in science, a subject area not specifically targeted by NCLB. Still, it seems reasonable to expect that if the law were actually improving math and reading performance, students might have an easier time with science as a result. As it happens, our overall science score dropped from 499 in 2000 to 489 in 2006, and our ranking fell from 14th out of 27 countries to 21st out of 30 countries. Scores also declined on both the PIRLS and PISA reading tests, but these drops were too small to be statistically significant.

Taken together, the results of NAEP, PISA, and PIRLS indicate that the No Child Left Behind Act has continued the legacy of failure established by its predecessors.

## Conclusion and Recommendations

Since the institution of ESEA and Head Start in 1965, federal K–12 programs have cost taxpayers roughly \$1.85 trillion in 2008 dollars. To get a feel for how large that number is, 1.85 trillion seconds equals 58,726 years—about 20 times the entire span of recorded human history. After four decades and nearly \$2 trillion dollars, after unsuccessfully cycling through an endless series of programs, we have enough evidence and experience to draw a solid conclusion: the federal government cannot significantly improve school performance.

One thing that the federal government could do that would undeniably help American families would be to stop taking vast sums of their money and funneling it into patently ineffective programs. By phasing out futile federal efforts in education, taxpayers would regain control of 70 billion of their hard-earned dollars every year. The effect of this financial windfall on American society would be significant: creating jobs, stimulating investment, and raising the overall standard of living.

It is well within the power of Congress to bring about these benefits. The first step would be to convert all existing federal K–12 education programs into block grants to the states. These grants should then be phased out completely over three years, giving states the time to reallocate their own personnel and resources. As the block grants are phased out, federal income tax rates should be proportionately reduced so that taxpayers retain the money that was previously being spent on ineffective federal programs. At the end of the three-year period, Americans would be enjoying a permanent \$70 billion annual tax cut.

The U.S. Constitution delegates the federal government no power to determine the content, methods, testing, or staffing procedures of American schools, but the Constitution is not the word of some supreme being. It is not *right* in some absolute sense, purely by virtue of *being* the Constitution. The wisdom of its exclusion of a federal role in education policymaking must ultimately stand or fall on the basis of the facts. If, as hypothesized in the Introduction, there were compelling evidence that federal government intervention in education would best serve the public's interests, then it would be foolish not to amend the Constitution to allow it. On the basis of the inconceivably expensive failure of federal interventions over more than half a century, it would be foolish not to cease them immediately.

### Suggested Readings

Arons, Stephen. *Short Route to Chaos: Conscience, Community, and the Re-Constitution of American Schooling*. Amherst, MA: University of Massachusetts Press, 1997.

- Boaz, David, ed. *Liberating Schools: Education in the Inner City*. Washington: Cato Institute, 1991.
- Coulson, Andrew J. *Market Education: The Unknown History*. New Brunswick, NJ: Transaction Books, 1999.
- Lieberman, Myron. *Public Education: An Autopsy*. Cambridge, MA: Harvard University Press, 1993.
- McCluskey, Neal. *Feds in the Classroom: How Big Government Corrupts, Cripples, and Compromises American Education*. Lanham, MD: Rowman & Littlefield, 2007.
- Schaeffer, Adam B. "The Public Education Tax Credit." Cato Institute Policy Analysis no. 605, December 5, 2007.
- Tooley, James. *E. G. West: Economic Liberalism and the Role of Government in Education*. London: Continuum, 2008.
- Walberg, Herbert J. *School Choice: The Findings*. Washington: Cato Institute, 2007.

—Prepared by Andrew J. Coulson

